COMPUTER APPLICATIONS
IN
MEDICINE

Publication Number 557

AMERICAN LECTURE SERIES®

A Monograph in

The BANNERSTONE DIVISION *of*

AMERICAN LECTURES IN LIVING CHEMISTRY

Edited by

I NEWTON KUGELMASS, M D , PH D , Sc D

Consultant to the Departments of Health and Hospitals

New York, New York

COMPUTER APPLICATIONS IN MEDICINE

By

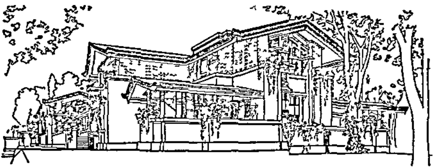
EDWARD E. MASON, M.D., Ph.D.

Professor, Department of Surgery,
College of Medicine
State University of Iowa
Iowa City, Iowa

and

WILLIAM G. BULGREN, M.S

Graduate Student in Mathematics
State University of Iowa
Iowa City, Iowa



CHARLES C THOMAS PUBLISHER
Springfield • Illinois • U.S.A

3274

Published and Distributed Throughout the World by
CHARLES C THOMAS • PUBLISHER
BANNERSTONE HOUSE
301-327 East Lawrence Avenue, Springfield, Illinois, U S A

This book is protected by copyright No
part of it may be reproduced in any manner
without written permission from the publisher

© *1964, by* CHARLES C THOMAS • PUBLISHER

Library of Congress Catalog Card Number 63-18522

With THOMAS BOOKS careful attention is given to all details of manufacturing and design It is the Publisher's desire to present books that are satisfactory as to their physical qualities and artistic possibilities and appropriate for their particular use THOMAS BOOKS will be true to those laws of quality that assure a good name and good will

Printed in the United States of America
S-3

To Progress in the Art and Practice of Medicine

FOREWORD

OUR Living Chemistry Series was conceived by Editor and Publisher to advance the newer knowledge of chemical medicine *in the cause of clinical practice The interdependence of chemistry* and medicine is so great that physicians are turning to chemistry, and chemists to medicine in order to understand the underlying basis of life processes in health and disease Once chemical truths, proofs and convictions become sound foundations for clinical phenomena, key hybrid investigators clarify the bewildering panorama of biochemical progress for application in everyday practice, stimulation of experimental research, and extension of postgraduate instruction Each of our monographs thus unravels the chemical mechanisms and clinical management of many diseases that have remained relatively static in the minds of medical men for three thousand years Our new Series is charged with the *nisus élan* of chemical wisdom, supreme in choice of international authors, optimal in standards of chemical scholarship, provocative in imagination for experimental research, comprehensive in discussions of scientific medicine, and authoritative in chemical perspective of human disorders

Dr Mason and Mr Bulgren present a fascinating treatise on the principles and practice of the newer analog and stored program digital computers in clinical medicine, medical research and hospital administration It is medically applied cybernetics, the new science that emerged from the mechanistic materialists of the eighteenth century and their modern counterparts The purpose of constructing a conceptual world in terms of machine analogues is to unravel the mechanism of life in health and disease The methods not only yield vital knowledge by a rapid process of piling up facts but cardinal wisdom in their simplification through effective integration of heterogeneous factors This great concurrence between learning and insight via electronic computers marks a new era in quantitative medical thinking that comes out of the qualitative advances of the past Yet the automation is as natural a growth as an oak emerging from current developments in biophysics, medical engineering and biomedical computing

The high speed electronic computer is a new tool, different from any ever created, for extending the scientist's mental activity in problem analysis, concept exploration and clinical diagnosis Small signals effect large outputs of medical information involving extensive data, multitudinal factors and mathematical manipulations Accumulated knowledge is stored electrically, correlated promptly and analyzed automatically by planned programming to provide prompt solutions to pressing problems without waiting and hoping The computer is a model of nineteenth century predestination, devised by Boole and Babbage, English mathematicians, in 1822 to relieve man of the drudgery of computing It provides all the basic theoretical and mechanical knowledge necessary for a blueprint of the modern giant machines, lacking only the thermionic tube to give it the speed of an electron instead of a piston This transcendent capacity to make reason sublime lay dormant over a century before American scientists began to create models of functioning living creatures in search of a *deus ex machina*

All computing automata fall into two great classes well represented in the central nervous system The digital, which counts electric pulses like beads on an abacus, and the analog, which measures continuous variables like scale lengths on a slide rule The former is more precise, the latter faster, but both must be programmed, i e , instructed to carry out the desired computation Digital computers developed on the principle of desk calculators perform mathematical computations heretofore unavailable because of the great time required for the calculations and the relative shortness of human life Significant is the remarkable ability of the computer to store information in the form of punched or magnetic tape as a permanent medical memory for the diagnosis and treatment of every human disease Such long-range development of a universal tool for the physician will accelerate the ultimate goal of clinical medicine Programming a problem requires several weeks, the calculations several minutes in comparison with hand computations that require several years Analog computers differ from digital systems as continuous functions differ from the discontinuous in mathematics They represent physical conditions by voltages and currents rather than by numerical means enabling satisfactory investigations of complex problems with many potential hypotheses via twenty operational amplifiers The authors' *experientia docet! Experto credite*

Natural systems are of such enormous complexity as to necessitate subdivision of the problem into elementary units The end-result of medical research on a system is to learn enough about it to describe it in mathematical equations and then resort to computers for the solution The human mind is a marvelous computer with 10^{23} computing elements but limited in speed and afflicted with forgetfulness, both corrected by man made computers of less flexibility and little creativity with 10^3 to 10^6 computing units The computers can only function in harmony with an active human mind, performing manipulations according to the instructions of the operator unless created with a prodigious 10^9 parts Electronic computers are thus endowing material forces with intellectual life without degrading human life into a material force Newer techniques of molecular miniaturization and current advances in computer organization may yet mime the higher mental processes Bergson considered intelligence in its original feature as the faculty of manufacturing artificial objects, especially tools to make tools, and of indefinitely stimulating the creation of increasingly complex machinery Chaucer thought that "A man hath sapiences thre, Memorie, Engin and Intellect also" But Goethe warned,

> *Am Ende haengen wir doch ab*
> *Von creaturen die wir machten*

I NEWTON KUGELMASS, M D , PH D , SC D , Editor

ACKNOWLEDGMENTS

THE authors owe a debt of thanks to E F Lindquist, J A VanAllen, the Administration of the State University of Iowa and to J P Dolch, the first director for the early and rapid development of a University Computer Center Research supported by U S Public Health Service Grant RG6500 brought the authors together in the Center and stimulated the computational research described in Chapter VI C C Spiker introduced the analysis of variance programs, compatible with *Design and Analysis of Experiments in Psychology and Education* which have been used in the research analyses D J Norton and H P Bechtoldt among others have contributed a great deal to the development of software in the Center as well as stimulus to the authors in more efficient use of the Center

H W Shipton read the section on electroencephalography and H E Hamilton the electrocardiography section of Chapter IV E F Mason, Professor Emeritus in the School of Journalism, and R T Tidrick, Chief of the Department of Surgery, both read the entire manuscript H P Bechtoldt read Chapter VI To all of these we extend our thanks for suggestions and corrections, but absolve of blame for any persistent deficiencies

There are many others at S U I and elsewhere whose contributions could be mentioned were there space Some have been referred to in relation to their published works There were many generous contributions of reprints and manuscripts, some of which had not been published prior to their receipt by the authors

Each of the authors would like to acknowledge the support and understanding of their friends and family and the assistance of many secretaries, students and laboratory workers whose multiple contributions made the research and book possible F Chernigoy's faithful work in the laboratory and Marilyn Sanner and Jean Dunlap's preparation of the final manuscript deserve special mention Also the assistance of C J Gochenour of University Statistical Service and his associates deserve recognition for their efforts in the development of a key-word in context bibliography, part of which was used for the book

Lastly, I Newton Kugelmass, Editor of this series deserves the credit for getting us started on such an enterprise as this book in the first place The publishers have been patient but responsive and efficient when their responsibility began and the product of their art speaks the rest of our appreciation This has been a more cooperative effort than most We have received from many and return our report to those who find an interest or need

E E M
W G B

CONTENTS

Page

Foreword vii

Acknowledgments xi

Chapter

I COMPUTERS ARE TOOLS 3

Computers do routine tasks and release man to think and communicate with his fellow men—What is a computer and what routine tasks do they accomplish?—Stored program and automatic decisions—Flow charts—Limits of data collection and analysis—Need for prospective studies, well defined problems and use and knowledge of standard experimental design and analytical methods

II HOSPITAL ADMINISTRATION AND OPERATIONS RESEARCH 20

The problems of management require vast files of information which must be used for decisions about day-to-day problems of supply and demand, traffic and reporting as well as long-term planning—Complete automation is under study—Even small hospitals have available now subscription services for some needs—Business procedures are most easily automated

III DIAGNOSIS 37

Physicians use a vast amount of information in their decisions—Man's forte is not memory but synthesizing thought—Machines can store data and make routine decisions—Physicians must program the machines and validate the results

IV ELECTROENCEPHALOGRAPHY AND ELECTROCARDIOGRAPHY 50

EEG research is statistical and deals with infinite data—ECG analysis illustrates curve analysis and leads to the question of information theory—Both have been greatly accelerated if not transformed to a new level of investigation by computer application

Chapter *Page*

V MEDICAL RESEARCH 70

Volume of data, complexity, speed required for analysis are all reasons for computer use in research—Varying degrees of computer use are illustrated from routine data sorting to doing the research in the computer by simulation of a biologic system

VI A WORKED EXAMPLE 84

An investigation of enzyme loss by the kidney is used to illustrate the use of a computer center and routine analysis of variance, multiple regression analysis, covariance and special programs for transformation—Writing a program in FORTRAN is illustrated—Problems of design of experiments and interpretation of results are emphasized

VII INFORMATION ACCESS 119

Physician focus of information flow—Creativity—Coding, storage, retrieval—Immediate access to the world's knowledge at your desk—Levels of retrieval—Defining the area of interest—Key words—Information theory, science and specialists—A new dimension of insight

VIII EDUCATION IN PROBLEM SOLVING 129

Labor saving machines started the industrial revolution—Time-saving machines now herald the automation revolution—More time for art of medicine promised plus time for study—Mathematics and logic in medicine—Self education and changes in medical education—*Physicians face new* problems from automation in our society—Education and planning required

References 135

Author Index 167

Subject Index 169

COMPUTER APPLICATIONS
IN
MEDICINE

Chapter I

COMPUTERS ARE TOOLS

AUTOMATION in industry and the use of computers in hospital administration and in medical research have stimulated a great deal of interest and perhaps some unjustified anxiety about automation and its effect on medical practice Installation and use of equipment is gaining momentum in hospitals, medical schools and research laboratories This book will attempt to outline some of the problems that exist, some of the things that have been accomplished and some of the current trends toward more universally applicable solutions For those who still have doubts about the ultimate use of automatic data processing in medicine or who fear the effect of these techniques on the practice of medicine it is hoped that the following discussion will show the potential savings in time, increase in efficiency and a new type of freedom which these new tools offer

The modern physician is so busy doing routine things which a machine could do that he does not at present have enough time for the personal attention that his patients used to receive Valuable physiologic and biochemical laboratory studies can be automated in such a way that considerable analysis and data reduction takes place and information is retrieved by the physician in much more useful and understandable form Automatic data processing techniques can provide part of the answer to many of the complex and challenging problems in medicine

The techniques will require hard thinking about objectives, how man functions and how the many routine activities could be automated This means a new order of magnitude of learning and a creation of jobs with new purposes, titles and perhaps new departments in our medical schools There are problems in medicine that may require special techniques but most medical problems are made up of the same type of questions that are asked in other disciplines Problem solving is an activity of mankind and physicians function in much the same way that other men function

What is a computer? The builders of computers have a much different answer to this question than the users of computers Occasional clinicians, because of their background or interest, may become well versed in the electronic details of how computers function and may be able to contribute to the further development of equipment which will have special advantages to applications in clinical problems Such crossing-over between disciplines is to be encouraged For the majority of clinicians, medical investigators and even basic research workers, the computer must remain as a black box The use of an automobile is an analogy familiar to all The driver may derive almost as great a benefit from knowing only how to start, stop and guide the machine as does the garage man who also takes care of it or the engineer who also designs it

At times a computer definition must include some information as to how it is being used For example, a pencil would not be called a computer if it were used in its usual fashion for making marks on paper A pencil might be considered an analog computer, however, when it is held up at arm's length by an artist in order to estimate the various dimensions of the landscape so that they might be transferred to paper in their proper proportions Such use of a pencil as an analog computer would not be considered misuse of the pencil since it is efficient and appropriate to use the pencil first as an analog and then in drawing the picture If primitive man were given a box of pencils with no information as to what they were or how they could be used, he might use these as tally sticks, taking one out of the box each time one of his "domesticated" animals passed into the fields in the morning and replacing one pencil in the box as each animal returned in the evening Used in this way, the box of pencils would become a digital computer and since stones or crude pieces of wood could be used as well, it is rather obvious that the box of pencils used in this way would be inappropriate and extravagant

The slide rule is an analog computer which can be used for multiplication, division, and even more complicated manipulations The well versed user may set up an equation containing a number of manipulations and arrange them in such a way that the information is maintained on the slide rule until the final answer is obtained With more complicated problems, it is necessary to incorporate a pencil-and-paper recording of the partial results and to work back and forth between the slide rule (which can be considered the arithmetic unit) and the pencil and paper which can be considered storage, part of input-output, and control If greater accuracy is desired, but mechanical assistance seems desirable, the next step in complexity would be to use a desk adding machine or a desk calculator Here again, part of the solution may be maintained within the calculator between various steps of the problem, but the paper and pencil are still often needed

The research worker who is analyzing curves and desirous of integrating might add to his equipment a simple planimeter which is also a computer and capable of determining the area within any regular or irregular curvilinear space If a planimeter is not available and a number of curves are to be compared, the worker may draw the curves on standard graph paper, cut them out, and weigh the cut-out portions of the curves in order to obtain an integration of the area beneath each curve The weight of the paper is a function of the integrated curve Each age has had its mathematicians and its makers of mathematical machines The modern computer is a relatively recent development

The modern analog computer is an expansion of original computational schemes such as the slide rule or the planimeter but is capable of much greater variety of uses and far more complex problems It is more versatile by virtue of electronic parts instead of mechanical parts The modern digital computer has shown the greatest development and the greatest promise of all computing machines because of the concept of a stored program

The entire telephone system is a big digital computer with which, by successive dialings, we eliminate large groups of phones until, with the last dialed number, a single address is selected and we then listen to a restricted segment of information, a window if you please, which limits the available signal The signal obtained may be a recorded human voice, programmed to play in proper relationship to real time so that the interrogator may learn the time of day By having the necessary local equipment and a phone book we participate in a system much more complex and expensive than we as individuals require or could afford Those subscribers with larger needs may buy more service in the computing system or more and greater variety of peripheral equipment or both

A modern stored program digital computer makes use of the telephonic technique of address points and stored information at these points The stored information may be parts of tables, data that is as yet unprocessed, statistical data that has resulted from partial or complete analysis or instructional material that is part of the stored computer program The computer is a step away from telephoning a machine to learn the time of day in that the computer is programmed to talk to itself, shift data around, work with it in a prescribed manner and, at various points along the way, to report out intermediate results until at the end of the program final answers are ready for our perusal

The analog computer, as its name indicates, uses an electrical signal such as voltage which is measured as an analog of some physical or measurable quantity under study For addition and multiplication a circuit is designed with adding units which feed in the appropriate additional amount of voltage or an amplifier which multiplies the voltage by the required amount Capacitors are used to accumulate the voltage over time as an analog of integration A capacitor is analogous to cutting out and weighing a graph of a concentration/time curve If several equations are to be solved simultaneously, the circuits can be so arranged

Such equipment is useful for small to moderate sized problems The equipment is relatively low in cost and fairly easy to assemble Rearrangement of the parts can be carried out with a wiring plugboard Coefficients in the equations can be altered by the use of variable condensers The equipment can be made available right in the laboratory If the research worker wishes to analyze a curve without mathematically deriving the equations involved, he may be able to accomplish this by manipulation of the dials on an analog computer and obtain a good enough fit for most purposes in biological work The analog computer is useful *on line* in the experimental procedure if the experiment has to do with electrical impulses, and at times a partial analysis may be used to provide an appropriate stimulus as a part of the experiment

The analog computer can be of tremendous help in teaching the applications of mathematics in biology as it is in other fields The analog computer may be designed for on-line simulation as a model of the experiment so that repeated adjustments can be made in the computer until the results correspond to what is observed in the experimental subject Thus, the accuracy of the analog which is under investigation in relation to the experimental subject can be rapidly checked and modified Such an arrangement is very useful in the development of hypotheses as to how physiologic control mechanisms seem to function in the body It is possible in this manner to investigate moderately complex problems where dozens of minor hypotheses must be formulated and tested Such testing by adjustment of an analog computer may be more economical than modifying programs for a digital computer

The digital computer can do practically anything that the analog computer can do It has been said that if problems require a large analog computer, it may often be more economical to simulate the large analog computer on a stored program digital computer than to actually construct the analog computer Problems which can initially be solved satisfactorily with an analog computer may grow in numbers of simultaneous equations requiring solution until it becomes necessary to transfer the work to a digital computer

The analog computer has a limited accuracy whereas the digital computer can be programmed to develop answers to any desired number of decimal places It is difficult to classify analog and digital computers as to the types of problems for which they should be used It would seem, however, that the digital computer is much more versatile and can be used not only for arithmetic problem solving, but is easier to program for the solution of problems involving many logical decisions This latter type of problem is more common in clinical research where the desire is to store and retrieve data, to rearrange data in different categories, and to study relationships of the incidence of signs and symptoms in relation to certain diagnoses The digital computer may make it easier to use a combination of logical and mathematical decisions

The chief advantage and advancement of the modern digital computer rests in its facility for storing the entire program within the computer Stored program digital computers have a number of black boxes designed for input, output, control of the operations and storage of data and reference material with varying degrees of availability There are two equally important divisions of computer equipment into *hardware* (or the actual physical computer equipment) and so-called *software* (the programs which are to be stored in the computer and which instruct the hardware in the sequence of steps required for solution of a given problem)

The stored program is made up of a series of *words* These *words* are in numerical form or so-called *machine language* Each *word* includes an instruction as to the step which is to be carried out, such as addition or subtraction, or some logical step such as determining which line of further logic is to be followed The *word* also includes an address where the information is stored which is to be acted upon at this step in the program and finally a second address to which the machine is to turn for the next instruction Thus, one instruction follows another

Data are read in from the input section or brought in from various storage areas, acted upon, rearranged, re-stored in other addresses for later use in the sequence of analysis Comparisons are made, tables are searched in the memory, long series of similar numbers are summed and perhaps arithmetic means and standard deviations and other moments of higher order are calculated and stored

Finally the results are fed out onto magnetic tape, or punched cards to be printed, or graphed by ancillary equipment, for inspection by the research worker It is obvious that to make use of the computer, the investigator must know what he wishes computed and must have a very detailed knowledge of the steps which are required to carry out this computation by hand, even though the actual labor of such an analysis might make the hand analysis impossible

The first digital computer, consisting of the digits on the hands and feet, predisposed man to the use of a decimal system of counting Digital computers actually make use of a binary system in which the decimal system can be simulated Even the digital computer is made up of electronic components such as transistors, relays, vacuum tubes, Ferrite memory cores, and other electrical components which have basically two stable states

It takes about three and one-third times as many characters to represent a given magnitude in the binary number system as compared with the decimal system Letters can be coded and manipulated in the digital computer so that the research worker may make use of Boolean algebra or some other system of symbolic logic The versatility of the digital computer is thus limited solely by the ingenuity of the user in defining his problem and in arranging the detailed steps or larger components of library programs and subroutines for the solution of his particular problem

The power of the tool derives from the accumulation of software as well as from the speed, memory capacity, and versatility of the physical equipment The tremendous potentialities are very seldom appropriately matched by the problems presented This is inevitable The equipment and software are the product of thousands of experts using the accumulated knowledge of generations of the most brilliant problem solvers The chief difficulty today is in formulating problems worthy of such equipment and restricting computer use to those problems that really need so much automatic data processing

Most physicians will have specific problems of a more limited scope than general problem solving, diagnosis and literature retrieval The requirements of efficient computer use are (1) the presence and recognition of the problem, (2) an approach to solution of the problem including certain stated hypotheses and methods of testing these hypotheses, and (3) translation of the testing procedure into machine language All of man's activity involves sampling and data reduction Our senses record only a very limited amount of the available data around us The temptation when confronted with something new and exciting is to overindulge

It is necessary, first, to recognize the existence of a problem which could be solved better by machine methods than by the direct application of pencil and paper and the inappropriate use of human brainpower Computers do not replace thinking They demand deeper thought They permit the thoughtful detailed sequence of logical steps used in solution of a problem to be used repeatedly without repeating the same human expenditure of effort Many different problems require parts of the same thinking Problem solving has repetitious aspects at all levels and this is what allows the use of programmed machines

Problems range from those of how to get from A to B (so-called applied research, or development, or just simple business procedure) to those problems concerned with defining A and/or B (basic research) The *problem* may be simply one of facilitation of some activity now carried out by hand which could be automated in order to save time and money or in order to continue the service in view of its expanding proportions Thus, an administrator may wish to automate a detailed billing procedure which will inform each recipient of the services and materials which he has received An electrocardiologist may wish to provide an ECG report of a very complicated mathematical and statistical analysis which is of known reliability and which would require hours or days if carried out by hand A psychologist may wish to provide hundreds of reports on standardized personality and intelligence tests so that most patients seen in an institution would have at least these basic test results for use by the medical staff An internist might wish that all patients seen in his diagnostic clinic have an array of standard questions answered for machine analysis before the patient is actually interviewed by his physician This might take the place of much of the so called "systemic review' and might allow faster and more complete review of items which are not immediately pertinent to the chief complaint and which might otherwise be neglected in some patients These are all questions of getting from A to B They are relatively simple to program

The second broad category of problems is the use of automatic data processing to help find out something about A and B These are questions that come up in the course of research and investigation at all levels To be practical, it is necessary to link such questions closely to their proposed means of solution It does no good to ask a question that has no detectable or potential method of solution. A problem must be defined in sufficiently clear and concise terms so that hypotheses may be developed and appropriate observations and manipulations carried out in order to seek solution. This is the most difficult area at present in the application of automatic data processing in medicine. This is where much heat is generated. This is where the art of medicine is championed and the invasion of machines decried. This is actually the point at which the living, self-reproducing computer finds he must "fish, cut bait, or go ashore." This is where the physician finds that he needs to know more about his problem and problem solving, including various mathematical and statistical techniques. This is where the statistician, if he wishes to work in the field of medical application, needs to know much more about medical problems. This is where the mathematician who wishes to help must be practical and attempt to apply his theoretical knowledge to the problem at hand. This is where it becomes evident that each man is dependent upon his fellow men and that they, in turn, are dependent upon him. Problems under such close scrutiny often become very personal and bring out the deficiencies of each and every one of us but also offer the challenges that mankind seems to live for.

Having "staked out" a problem which seems amenable to automatic data processing, one or more hypotheses must be stated. An hypothesis is a statement of certain relationships which must then be tested by experimentation and analysis of the resulting data. There may be many separate hypotheses about a problem, since each of these is usually concerned with some detail. For example, the problem might be one of how to diagnose more accurately and earlier and more easily a group of related diseases. The specific hypotheses, then, would be concerned with the relationships of signs and symptoms in these diseases to the diagnosis of certain sub-classifications of diseases or specific diseases. The testing of a hypothesis may actually reveal the need for an increased scope of the problem statement rather than an immediate attempt at solution.

As the rigor of the investigator's approach to the problem increases, as the questioner comes closer and closer to the basic facts of the problem with which he is working, there are usually recurring loops of thinking which refer back to the statement of the original subsidiary hypotheses or of the entire problem Often the worker finds, in the course of trying to state his problem, that he did not really understand the problem in the first place This is the real power of automatic data processing and a theme which deserves some repetition—namely, that machine methods of solving problems demand a level of thinking, an intensity of thinking, a strictness of semantic activity which is often a revelation to the computer neophyte but which is necessary before any real progress can be made with machine use

There is a continual reworking of the questions, going from very general statements to statements that are more and more specific This is assisted by flow charting the actual steps in the solution of the problem A flow chart consists of a series of diagrammatic blocks, each of which contains a statement as to what step is to be taken or what observation made, or what set of calculations carried out before the next succeeding step can be taken

The initial flow chart may consist of only two or three blocks in the diagram According to the problem, each of these blocks must then be worked out in greater detail until, finally, a series of specific logical and/or mathematical equations can be written In the early days of computer use, it was necessary to carry such flow charting and programming all the way to machine language This is no longer necessary since there are now available special computer translation programs which allow the programer or research worker to write out his proposed program in an understandable sentence structure, using almost normal English words Providing certain rules are followed, the computer programed with a translation program will then convert the recognizable human statements into numerical machine language In addition there are available many sub routines so that if the worker, for example, wishes to obtain a square root of a number, he can simply incorporate "the square root sub routine' with all of its machine language at the specified point in the analysis Many powerful sub routines have been incorporated in larger routines which, in turn, may carry out a whole complicated statistical analysis at an appropriate place in the desired program Thus for example, the research worker may wish to determine the simple correlations between several variables and might incorporate the library program for correlation analysis into his program

For the experienced programer with a library of these subroutines and special programs, these powerful tools can be assembled simply by the appropriate writing out of a few brief statements It is thus possible to make use of the accumulated knowledge and skills of the most gifted mathematicians and statisticians of our time and of previous generations plus the knowledge and skills of thousands of electronic engineers and other specially trained people whose vast efforts have been called together and assembled in a specific program on a given computer to answer your question

The translation and assembly programs which have been written for various computers include many automatic methods of recognizing common errors in writing programs so that after the program is written in recognizable symbols and run into the computer for translation into machine language, there is an immediate return of questions from the programed computer calling attention to a variety of mistakes in the way in which the program is assembled

Many problems do not require automatic data processing Some problems could better be carried out with a set of hand-sort cards or possibly with punched cards to be sorted by machine Some problems are best suited to small analogue computers Others require a digital computer with a stored program Some problems have a great deal of data which requires efficient access to the computer and large storage areas but relatively little actual computing Other problems require very little data to be read into the computer, but require lengthy computations before the answers are printed In most instances, the actual computations are so fast that the limitation in efficiency depends upon the rate at which information can be read into the computer and read out again

A problem may have several alternative mathematical approaches which make it more economical to use the larger computer so that the various methods can be tried Then, after the best approach is found, a special purpose computer can be used for subsequent repetitive analyses using the approach already worked out with the more elaborate equipment In some instances computers must be available right in the laboratory where the research work is being done Some of these installations must be so called "on-line" so that computations may be carried out during the experiment and with some of the results used to, in turn, modify the experiment, a transponder function The actual equipment used for solution of specific problems will, of course, often depend upon what equipment is available to the research worker

The cost of data processing with various available types of equipment must be taken into consideration Someone may need to decide upon the relative value of the project in relation to other demands on the equipment There may be facilities of hardware or software available elsewhere which would make it more economical to do the data processing even outside of the original institution in spite of there being equipment locally available

With expanding facilities and expanding problems, it has become evident that a variety of installations will be necessary and that some of these could most efficiently be ranged in a hierarchal and compatible fashion so that small computers can be used in laboratories or hospitals up to a point and then, when larger facilities are needed, the problem can spill over into the next larger installation It is inevitable that some problems will require the use of large regional data processing centers and computers such as those which have been established for national defense There are expensive special purpose installations which should be available to investigators who would travel to the installation for a few weeks or months in order that such equipment not be wastefully installed for only brief requirements

A potential user of computers should take warning that it is not wise to attempt to store all data for all possible problems for an indefinitely long period of time in order that any future research worker with a brilliant idea may re-sift all of this data in the expectation of obtaining the specific information required to prove or disprove his hypotheses This approach is doomed to failure Every line has an infinite number of points It is possible, therefore, on the most minute problem to assemble more data than our best computing equipment can handle

It has been estimated that any problem which requires more than 10^{13} manipulations cannot be handled by present day computers Even this number of computations would require at least three years on our present fastest equipment The clinician would like to store all of the information that is on a patient's chart, but all of the information that is on a patient's chart is only a very small part of the data which are available about that patient So, even if he could store all data from all charts on all patients, this still would be an abstract

The most common approach to clinical research, the path of least resistance, has been to collect a large amount of data or as much data as can be handled and then to survey such data in all of their permutations, combinations and statistics to see if there are any striking relationships which might signify some new insight into the overall problem The general idea too often seems to be *to collect a lot of data and then to do a retrospective study picking* up hypotheses along the way This approach implies that the data will be collected without knowing what questions are to be asked Such an approach to be successful must theoretically include the greatest amount of data that can possibly be stored

There are certainly some key data, such as patient age and sex, which are worthwhile collecting in almost any clinical investigation and which will probably be necessary for retrospective studies if and when hypotheses are stated There is no reason why such retrospective studies should not occasionally be continued and should not occasionally be productive However, the commonsense approach and the scientific method requires an initial stating of some hypotheses and establishing a protocol for collection of *data in a prospective study Later on, newly formulated retro*spective studies may still be made with this data

At least, if one starts out in the prospective mode, there will be better collection of those data which are included in the protocol, and there is a better chance the data will be useful It is, of course, an oft' repeated truism that no information can be extracted from data that is not intrinsically present in the data when it is originally obtained as a sample It is equally true that when large amounts of data are analyzed from a good protocol that the techniques of analysis may bring out information that would not otherwise have been observed, just as a pH meter, spectrophotometer or electrocardiograph reveals relationships which could not be evaluated with the unaided senses

As soon as the investigator accepts the fact that he cannot study the universe and that most of what he studies is a relatively small sample, he may then begin to define that part of the universe in which he is interested He may wish then to define what a random sample would be for a selected portion of the universe Sampling will depend, to some extent, upon the type of experiment and also on the analytical techniques which are available for the analysis of the data Attention should be paid to various control conditions which should be met in order that the study is confined to the area of declared interest

A number of parameters are selected for measurement and perhaps other observations of a non-parametric nature [without a scale] are also recorded Environmental conditions should often be included The experimental design should be considered with great care and the most important dimensions formulated in such a way that they can be separated out in the final analysis The cost of the sample must be considered along with the size If the cost is too great, perhaps one dimension can be eliminated by restricting the experiment to a limited age group or a limited sex or certain limited conditions of the experiment The cost must be considered in relation to the desired accuracy of the results

The research worker is often interested in studying a particular cause-and-effect relationship and suppressing those interfering forms of random activity, error and noise which obscure the signal in which he is interested When the expected result, or the signal of interest, is not pronounced, it may be desirable to use repeated observations or repeated subjects under similar conditions in order to average out the interfering noise of sampling, analytical error, random variation, background noise, etc

The research worker who has a problem which allows him to sample a very precise curvilinear relationship with little detectable variation from experiment to experiment is fortunate He may then go directly to curve analysis and the interpretation of the observed relationships in terms of specific models In much of the medical research today, the observations are not so concise The data with which we work requires statistical treatment before we can detect the actual relationships

It is necessary for the worker to understand sufficiently statistical methods so that he may decide upon the risks that he wishes to take in calling a result significant when it really is not significant or, on the other hand, in saying that there are no significant differences between two groups of data when, in actuality, there may be a true difference These two types of error are always present Anything done to minimize one type of error automatically increases the other

Finally, the requirements for efficient use of automatic data processing *should at least include some vague idea before the* experiment *is undertaken as to the interpretation and possible* generalization which will be made when the experiment is finished *and the data is analyzed The best laid plans would allow the* research worker to write out his conclusions before the experiment *is undertaken and then, at the end, to simply pick up the discussion* which best applies to the results obtained Actually, our hypotheses *are seldom so precise and our ideas as to the model under study* are seldom so accurate as to predict the alternative results to this extent Even when the results do coincide with some anticipated formulation, the mere perusal of these results stimulates one's thinking to the extent that new conclusions and interpretations become evident that had not heretofore been considered

There is, or should be a great deal of education for the investigator at all stages of his work The hypothesis and the original problem may need to be revised at any stage of the program development, its analysis, or after obtaining the final results of the computer analysis Any research project, and the use of automatic data processing as well, can break down completely at any stage if the human component ceases to function (think) It is human nature to expend the greatest effort in getting the program to work and in obtaining the data for analysis, and then to relax (cease to think) when the print-out is obtained Computers are programmed to make decisions, to do routine thinking, but not truly original thinking

One of the first duties after obtaining the computer print-out is to check the results for their consistency, validity and possible hidden meaning The computer is a "dumb brute" which will turn out volumes of stupid mistakes if it is so programed, or if inaccurate and inappropriate data are fed into it The computer technologists have built in many checks and balances which take care of most of the possible errors that might occur as a result of the actual equipment The standard translation programs which have been written for translation of newly written sets of instructions to machine language programs and many of the library programs for special analyses have written into them many checks which will prevent computational and common logic errors from creeping into the analysis There is nothing in the programed computer, however, that will prevent the investigator from allowing errors to creep into his data or even worse, from attempting to solve a poorly formulated problem by an inappropriate method

Good programs print out not only the final results of the analyses, but also enough of the intermediate calculations and conclusions so that the investigator may check the reasonableness, consistency and validity of the over-all analysis It is helpful to have the automatic data processing equipment hooked up with appropriate ancillary equipment to print some of these intermediary results in graphic form or to actually take the printed sheets and graph the intermediary sets by hand It is possible thereby to pick out completely wild values and to go back to earlier steps in the program or to original data and find the errors that have been overlooked

Finally, if all is well, the actual final equations or statistical tables must be studied so that valid conclusions can be drawn, and appropriate action taken, depending upon whether the problem has been one of a business procedure, basic research, clinical research, or otherwise Some practical application of the result may then follow such as the sending of thousands of itemized hospital bills to patients, or a rearrangement of certain facilities in the hospital for optimum use If the program has been one for basic research the investigator may then decide if his hypothesis is supported or false and may proceed to develop new hypotheses about the problem or perhaps turn to some new problem

The secret to success rests in having a well formulated concise, answerable question in the first place Automatic data processing without this leads to expensive processed data and nothing more Large amounts of data can be boiled down into powerful answers only to the extent that many hours of logical thinking about a problem have led to a good computer program There is a direct relationship between time spent on statement of the problem and time saved in data reduction

Chapter II

HOSPITAL ADMINISTRATION AND OPERATIONS RESEARCH

THE focal point around which all hospital problems in automatic data processing should focus is the patient, who must be well cared for so that he will receive maximum medical benefit and who must be satisfied so that he will look upon the hospital as an efficient and attractive place to which he would be willing to return at any time that he is advised to have further treatment

The hospital administrator may look upon his job as one of regulating a large network of interdigitating traffic of patients, personnel and materials He must be prepared to satisfy demand without costly excesses of supply The problem is a continually changing one but to a certain extent he is able to predict the changes before they occur if he has information about the national, regional and local changes in medical practice and hospital care.

Each hospital has a certain obligation to contribute to this over-all fund of knowledge This is usually fairly easy in situations where the internal workings are well studied The administrator must be prepared to alter hospital facilities in relationship to the appearance of new special purpose hospitals in the area He must understand some of the epidemiology of diseases and the effects of aging of the population on the types of care that must be provided Innovations in treatment and changes in attitude of patients and physicians as well as third party influences such as insurance plans and governmental participation all have their effect on hospital usage

The administrator has a continuing problem in keeping track of people including physicians, nurses, dietitians, pharmacists, laboratory technicians, housekeeping personnel, carpenters, plumbers and electricians as well as inventories of expendable and permanent equipment, medical supplies, food, drugs, blood, fluids, linens, uniforms and other items, all of which require a considerable amount of record keeping, planning, scheduling and regulation

The nursing supervision of a large hospital is a multifactorial scheduling problem which could probably be programed or investigated within a computer as a management game There are many types of special nursing areas and many different types of training and experience represented in the nursing staff plus hours, illness, vacations, emergency intense care situations, etc , etc , etc All of this must mesh without mess

Architectural changes in the plant itself are continually required and can be planned for to the extent that trends are apparent Consider, for example, the effect on hospital usage that came about during the poliomyelitis epidemic of a few years ago, the facilities that ultimately were provided and the more recent diversion of these facilities to other purposes Consider what the possible effect may be of gastric freezing if it turns out to be an effective treatment for peptic ulcer What changes would occur in hospitals throughout the country if a major breakthrough came in cancer therapy?

When new services are added in one department this often influences the traffic of patients and demand for room in other areas Questions arise as to whether subsidiary satellite laboratories or X-ray machines or fluoroscopy units should be installed, or whether perhaps whole departments should be moved These are but a few of the problems or types of problems that the hospital administrator is faced with and which require the collection of large amounts of data and many man hours of tedious analysis if it is all done by hand

The hospital administrator also has obligations to the actual care of the patient It is obvious that efficient business procedures and billing must be provided so that patients know what they are billed for In the hospital the patient deserves an optimum opportunity to find his way through the out-patient department, the various consultations and special studies, a minimum of wait for admission and reasonable assurance that when major operative procedures or other treatments are scheduled that they will not be canceled for want of the right blood type or because of an inadequate supply of operating room nurses

There is much that the patient does not know about, which also affects the efficiency of his care related to the flow of information from biochemical, hematological, bacteriological and other laboratories, from x-ray, from consultations between physicians, from the electrocardiographic room All of this information flow lies in the province of hospital administration and is susceptible to modern automatic data processing techniques

The administrator also has a problem in running the record room and in tending to the coding of diagnoses, keeping track of operations, complications, deaths, autopsy records, admission and discharge data, records of internal transfers of patients between services, information about patients' home addresses and relatives as well as records of current hospital location of the patient and the physician in charge, data required for insurance forms and the summary reports to be used in administrative offices and by the various services to fulfill the requirements for hospital accreditation and approval of training programs

The administrator would like to have a proven packaged system of known cost to solve all of these problems He would like a hospital electronic brain which, like a nervous system, would be large enough and efficient enough to co-ordinate all of the complex data gathering parts with the necessary means of appropriate presentation of information at stations throughout the hospital according to needs All of this requires a computer which should be attached to random access storage on the one hand and to larger computers to which it could spill some of the workload when needed Such an automated patient data system is theoretically possible today as far as the hardware, software and some peripheral equipment are concerned It needs electronic engineers, programers, operators and systems designers together with properly informed and motivated physicians, hospital administrators and medical research workers in order to make optimum use of such equipment with a consequent saving in cost to the patient, increased efficiency in patient care and use of the collected data for study of diseases and epidemiology

Such a data center would be available also for basic research and in education and could be used not only for the day-to-day feedback of information but also for special training of medical students and post-graduate students Networks of teaching machines could be attached to provide individualized programed instruction This would obviously be a complete solution to the problem presented above which the hospital administrator faces

Such a packaged system does not presently exist However, such a system is more than a dream and is being worked upon from many different directions There are somewhat similar problems in business that have been solved by computer analysis The large computer manufacturers and special organizations supplying consultation services have programs for solving scheduling problems and the problems of the best pathway through a complicated network with the least manpower and cost in material and with the greatest probability of reaching the desired destination Large businesses have played games with their business simulated on a computer to see the effect of changing procedures, offering new services, and changing the sizes and location of various components in the system These are not unlike the problems outlined for a hospital There are also similarities between the flow of traffic in a hospital and regulation of air traffic or the communication systems for national defense

It should not be surprising, therefore, to learn that the System Development Corporation is working on a patient data, computer based, system and that they have simulated a hospital within a computer using data obtained from actual patient records, programing the computer to carry out the various functions which would exist in a real hospital At the present time general input and output devices are used such as are found in standard computer installations

The ultimate arrangement of peripheral equipment for collection and display of data will depend upon the results of studies being carried out in other places It is possible that the physician would view a text stored in the central system through a peripherally located cathode ray tube and that the display might be a printed text or digital or graphic display depending upon the information requested

For study of administrative problems in the simulated hospital six wards of ten beds each, a pharmacy and two laboratories were established in the computer For the study of problems related to patient care three of the wards were provided with a great deal of information from actual charts, a primary diagnosis and up to three secondary diagnoses All of the necessary data were prepared prior to the experiment and placed on punch cards so that the information could be entered into the computer As the game was played additional information and requests were entered into the computer with the typewriter and the output came as typed or printed information

Much of the experiment was centered around the physician's *order By having the physician's order immediately stored in the* computer and all foreseeable consequences of this order of a clerical nature automatically taken care of, the nurses, pharmacists, business office and other involved personnel would have no further clerical work relative to the physician's order It is part of the game to have information reported either by hospital personnel or by monitors or both relative to vital signs according to the requests of the physician in charge The computer automatically would receive and maintain this data in its appropriate form and would print out on request these records for any given patient A certain amount of necessary feedback is arranged for in the program so that the nurse not only knows when an order has been made by a physician but knows that the pharmacy has received the order Also when the nurses are reminded that it is time to give a drug they report back to the computer when the drug has been given The computer program also keeps a running inventory of all supplies in the hospital and at predetermined levels of depletion, drugs and other supplies are ordered

When the physician orders a laboratory test the results are automatically consolidated into the patient's record in the computer The nurse is provided with a schedule of all the expected blood letting for that morning's laboratory tests The laboratory is kept aware of the expected workload For those laboratory tests requiring special feeding arrangements the computer notifies both the nursing station and the diet kitchen Even the discharge summary is automatically taken care of This then is the simulation approach and a computer is being used to seek the computer based solution

Another approach to finding solutions for the many problems of a complete data processing system is to set up regional medical centers and have them go ahead with equipment now available, work out techniques, integrate new equipment as it becomes available and actually seek information as to what is needed in both hardware and programs The first of these has been set up at Tulane University Medical School They are working out methods using special machine mark sense cards which can be marked by the physician, nurse or other person accumulating data for the patient's record These cards are then run through a reproducer which converts them to punch cards and the data can then be used in a routine manner with a card reader or converted to magnetic tape or disk storage They are working out means of recording all types of data that is present on a patient's chart including the history, physical examination, x-ray examination, laboratory tests and diagnosis codes A wide variety of computer applications are under study and the principal objective of this study is to seek means of application of automatic data processing to the types of data storage and retrieval that would be of greatest interest to the general practitioner and specialists who spend much of their time in diagnosis, treatment and routine patient care There are other large centers where special applications are under study in relationship to neuropsychiatric, neurosurgical and neurophysiologic problems

There is a study in Baltimore of central processing of accounting for some six hospitals which, if successful, may provide a nonprofit data processing service bureau for all of the hospitals in the state There are many university medical school computer facilities set up by departments within the medical school or in co operation with electrical engineering departments or the university at large

At the State University of Iowa Hospital patient billing applications have been a relatively recent exploitation of facilities that were originally started in the 1920's for university and hospital general accounting with card-sorting machines The University

Statistical Service Unit has handled all of the university general accounting and registration for some 12,000 students plus a great deal of statistical work for nearly all of the departments and schools in the university Recently the card sorting machines have been supplemented with a magnetic tape computer

In anticipation of this addition members of hospital administration, Statistical Service and the computer company spent some 18 months meeting together, defining all of the most detailed aspects of the problem and working out programs for carrying out hospital patient billing procedures A master tape was prepared which contains the registration biographical data and hospital numbers of some 500,000 patients treated at the University Hospital since its inception in 1927 Every out patient is now included in this identification system as well as all of the patients admitted to the hospital All hospital charges are put on punch cards and entered into a computer program which places them on the appropriate patient account on a magnetic tape Each in patient account is updated and printed out for business office use on each of the five work days during the week At the time of an in patient or out patient discharge, the discharged patient account is assigned an index number initially set at (9) and reduced by (1) on the occasion of each subsequent daily updating At the end of nine working days after discharge, the account having accumulated all "late" charges, is automatically printed out and made available to the hospital business office for patient billing or insurance submission, as the case may be

As bills are paid receipt information is also added on punch cards and when a bill is paid in full the computer prints out a statement to that effect and the account is removed from the tape The Iowa Medical Service Plan which acts as an accounting and collecting agent for private staff physician charges also uses a similar system In addition over the years many special medical studies requiring data sorting and simple statistical analysis have been carried out in the Statistical Service

In 1958, a separate University Computer Center was also established Originally this was a small open shop center but it has grown to a larger closed shop center with operators, programers and higher level personnel to help individuals from the medical school and the entire university who have computational problems that they wish to run

Much of the support of this center has come from its two largest users the Iowa Testing Program under Lindquist and the physics department satellite research under VanAllen Coincident with the growth of the Statistical Service and the University Computer Center there have also appeared many small computers about the campus of both analog and digital type Many of the analog computers are located in the medical school Some of these are so special in purpose that perhaps they should be classified in a category with standard electrocardiographic machines, which though analog computers, are committed to one narrow type of service function The Volemetron is an example of a commercially produced portable analog instrument programed for determining blood volumes with radioactive human serum albumin or with radioactive chromium 51 tagged red cells Most of these analog computers have been obtained for special research projects and some have been built in the school's section of Medical Electronics

It is inevitable that with the growth of data processing in a medical school and in a university that some users will desire equipment much larger than they alone can afford and that there will be continual pressure to increase the potential of the central computing center At the same time individual users will be faced with the decision as to whether they would prefer to do their computing in the large center and pool resources with other users or seek the advantages of having a smaller computer in the immediate area of work

There is also a continuing problem of personnel Computer science is a new interdisciplinary field drawing upon mathematics, electrical engineering, psychology and specialists from almost any field in which an individual becomes sufficiently interested in computer science to wish to devote his major effort in this area As these individuals become more involved in computer science they may be looked upon by their original specialty area as somewhat contaminated rather than pure academicians and scientists No apologies need to be made for computer science Computer facilities have risen out of a need and are looked upon too much as service centers, but there is a great deal of basic investigation involved in solution of the many problems These centers should be looked upon as a focal point of problem solving, theory and research and the problems posed are as varied as the nature of man Man's greatest challenge ultimately is in generalizing problem solving and the best problem solvers amongst us should be found and encouraged to work with or in such centers There will inevitably need to be large central computing facilities available in universities as well as smaller computers close to the areas of heavy routine workload such as in a college of medicine Probably the ultimate solution will be peripheral satellite equipment which can be used for smaller routine problems and a large central facility which can be shared with other users by special programing and coaxial cable

The administrator in nonuniversity hospitals may have a simpler problem in deciding about equipment and computer applications since the immediate applications will be recognized as those of business procedures for which there is excellent and fairly reasonable equipment available Hospital accounting programs have been worked out and would need only to be modified for local adaptation There is now available random access equipment for storage of data which would make it possible to expand the business procedures to include the handling of laboratory data and some of the data which might appear on a good discharge summary for each patient If the basic equipment for business procedures can be justified, the same equipment is then also available for clinical research It cannot be overemphasized that expert advice and extensive co-operative planning of objectives, equipment and program must precede equipment acquisition to succeed

Perhaps the easiest and most certain way of obtaining immediate benefits of modern data processing techniques is to participate in national studies of defined scope such as those of the paranatal study or a broader based study such as PAS One of the best nation-wide studies thus far is that of the Commission on Professional Hospital Activities, Incorporated which was established in 1955 as a nonprofit corporation to take over the Professional

Activities Study (PAS) As of January 1, 1962 the staff included seventy five persons variously trained in medicine, biostatistics, data processing, nosology, records and administration This group is conducting research on hospital clinical data needs and methods of collecting and handling such data Their files contain over six million clinical record abstracts Analyses are made for reporting back to the participating hospitals and a number of special studies have been conducted relative to morbidity, mortality and medical and hospital practices

The second major program in this study is the Medical Audit Program (MAP) which now is being used in thirty eight hospitals with a discharging rate of approximately 400 thousand patients per year This latter program allows the medical staffs of the participating hospitals to evaluate the quality of the medical care being provided and to look for special problems

The PAS abstracting service is subscribed to by some 230 hospitals in thirty two states, Washington, D C , Puerto Rico and three of the Canadian provinces Growth is continuing at an accelerating rate with over fifty hospitals added per year now compared with about twenty five only five years ago By January 1963, over ten per cent of current patients discharged from general hospitals in the United States were included in this sample The entire study is planned around case abstracts Case abstracts are produced by nonprofessional but specially instructed clerical help in each participating hospital

A single case abstract takes about five minutes to prepare These are sent to the central office for processing and a charge of thirty cents is made to the hospital for each abstract submitted This charge provides about 80 per cent of the cost of the entire project The additional expense has been made up from various grants and monies from research projects For the participating hospital, this plan provides an automatic means of evaluating the quality of care without increasing the workload of the physicians

A physician, Virgil N Slee, has been the director during the first decade of the study and continues to guide the service He stresses certain important points which could well be adopted in other large studies and might serve to answer some of the strongest *criticisms of automatic data processing in this* general area Data processing means data reduction and the development of statistics, so that the important individual may be lost from view In the PAS study, however, *the patient is retained as not only a fundamental* unit of the study, but if he is unusual in any observed respects, this information is carried through to the final reports so that the interested physician, investigator, or medical audit committee may make use of this most valuable bit of information

This does not mean that statistical studies are not also carried out The study is well designed and monitored for uniformity and accuracy of data A surprisingly large amount of information is obtained for use in study of diseases, their incidence, treatment and complications plus data which is useful for study of hospital utilization and local hospital practices Over the years the permanent record of patient abstracts is added to so that studies can be made of trends which should allow prediction of new hospital needs and areas where economies could be introduced as certain types of activity are eliminated

There are a variety of monthly and semi-annual reports which are automatically sent back to each participating hospital for use of individual physicians, department heads, hospital committees and administrators

Such a study should be a stimulating and beneficial influence on medical practice in general The legal yardstick for quality of medical care has, until now, been the current practice in a particular community But, hospitals which elect to participate in a nation-wide survey of this sort are, in a way, declaring that they belong to a nation-wide community and that they wish to *compare certain aspects of their medical practice with the national* average Comparisons between hospitals have been made as to the distribution of length of stay for patients according to selected admission diagnoses Thus, certain hospitals and certain physicians have been identified in a confidential way If a physician is extravagant in admitting many insured patients simply for diagnostic purposes, this becomes evident If excessive hospitalization is used for patients with tonsil and adenoid operations or some other minor ailment or operative procedure, it is evident in the automatic reports

All reports are coded in such a way that a physician or a hospital remains a statistic only except to the people directly involved in that hospital Of course, if a group of physicians wish to publish the results of their participation in this study, they may do so One hospital in which the laboratory hemoglobin values were abnormally low studied the reasons The effect of correction of certain aspects of the laboratory procedure resulted in a significant rise in the hemoglobin levels being reported Along with this change, a study was made of the number of blood transfusions being used in the hospital and there was a very remarkable saving in numbers of transfusions This, of course, resulted in an economy in time, space and personnel and also must be recognized as a worthwhile reduction in risk to the patients, since many patients must have been receiving blood transfusions who did not really need them

There are interesting ways of checking poor medical practices such as, for example, observing the number of patients discharged with sugar reported in the admission urine analysis, but with neither a blood sugar analysis nor a coded diagnosis of diabetes Studies have been made of the use of antibiotics in different hospitals, and over a number of years, together with a study of the relationship of antibiotic use to the use of the bacteriology laboratory Review of data from seventy six hospitals comparing the ratio of antibiotograms to antibiotic usage in 1959 and 1961, for example, showed a 31 3 per cent over-all increase in this use of the laboratory In 1961, some 26 4 per cent of patients given antibiotics were also provided with specific sensitivity studies For hospitals with under 5000 discharges per year, the ratio in 1961 was 0 104 The highest ratio, 0 310, was observed in hospitals classified as medium large with 10,000 to 14,999 discharges per year

Hospital staff sponsored quality control is generally accepted as important to the maintenance of a grade A hospital To do this effectively would require an unrealistic amount of physician time without setting up certain rules and limits for automatic screening of the bulk of records Automatic data processing permits common base problems to be assembled so that they can be recognized Patterns show up in certain areas and become self-evident as a result of the various types of reports and the permutations and recombinations of data that have been used to bring out such problems if they do, in fact, exist

Although the primary purpose of PAS is to provide information of immediate use to the participating physicians and hospitals, the abstracted information is permanently stored and available for large scale statistical reviews which should be invaluable in epidemiological studies on a nation-wide scale It is a study which the smallest hospital may participate in without any investment in equipment and without duplicating positions for any or all of the highly trained personnel required to develop and run such a program

Such a remote abstracting service, of course, does not solve the day-to-day data processing problems, many of which need to be on-line, and which the ultimate and optimum system should include There is also a great likelihood that larger hospitals which have some access to computer equipment may mistakenly feel that an external program such as PAS is an unnecessary duplication of what could be carried on locally One total solution cannot at present be applied to all medical data processing problems Some problems require a small computer installed on the spot, other problems can be best handled in a larger computer center within an institution, and still other problems can most economically be carried out by some larger outside agency

With present-day competition for time, money and equipment, it may seem illogical to pay an outside agency to do computing when a large computer *is available in the institution, but it would* seem to be much more economical to use PAS or other outside services and to reserve the use of local computer facilities for other studies which cannot be as efficiently carried on outside of the institution Information can be abstracted from their national storehouse of data by PAS and sent to participating institutions in the form of reels of magnetic tape so that any desired special statistical studies can be carried out with use of local computing equipment

Another important advantage of PAS is the institution of an optimum coding system which is standardized and acceptable to the majority of physicians There is a great tendency for hospitals and departments to destroy the effectiveness of a coding system by splitting off parts of it for redefinition This is always done in order to make the system better for a specific and immediate interest, but these interests change and soon the system is so altered that case histories cannot be recovered by diagnosis No system can be a good general system and at the same time very specific in narrow areas of interest When there is continual disagreement between physicians or departments about the coding it is often easiest to get outside help Then those who dislike the system because it won't do their research for them can blame the outsider instead of each other A good general coding system used by many hospitals must be used and maintained in a pure state If special studies require additional coding this is easy to arrange but it must be an addition to and not an alteration of the standard code

Another example of a study of national scope is the American Medical Research Foundation Paranatal Study which is a data collection program designed to improve the medical care given to mothers and babies Again under the direction of a physician, Sydney S Kane, this program has gradually evolved to a system which tabulates information from 145 hospitals in forty-three different states, District of Columbia and Puerto Rico The data is collected at one central headquarters, computer processed and then distributed back to the individual hospitals for their further analysis The data is sent back in three different forms One set of tables represents the individual hospital Another set concerns the regional area in which the hospital is located A third set of tabulations makes comparisons for the nation as a whole This program is also built around abstracts of individual patient records Although the original study was designed to improve neonatal and maternal care, there has resulted some interesting information of use to hospital administration and planning It is current practice in some areas in the country to keep both mother and child in the hospital for a much longer period of time If further studies reveal that no difference exists in maternal or infant mortality or morbidity by virtue of the longer stay, then a great saving in hospital bed utilization might be made

Observations also have been made which indicate an aging of the obstetrical community The progressive rise in multiparous mothers and a decrease in primiparous mothers indicates a long-range decrease in the requirement for maternal beds This, of course, is probably balanced by the increasing demand for geriatric facilities, but the trend generally and the trend in specific hospitals can be identified and could be used in planning of the requirements for the total hospital need and the most efficient distribution of facilities within a hospital

The American Medical Association, under the stimulus of this type of study, has, through its committees, looked into the legal aspects of this type of data collection A number of states have adopted legislation, but the Nebraska Act is cited by the Committee on Maternal and Child Care of the Council on Medical Service of the American Medical Association as an example which other states might follow in enacting such legislation as it is needed The Nebraska statute protects the confidential patient-doctor relationship but makes possible the reporting of information about patients as long as it is kept confidential (codified) so that committee work in hospitals, health surveys and other scientific investigations can be carried out The Legislative Bill 326 from the Seventy-second session of the legislature of Nebraska is considered a model bill *

There is increasing interest in operations research or study of hospital costs and planning of ways in which automation might reduce expenses and increase the efficiency and productivity of a hospital Reviews of some of these studies have been made by Flagle and also by Blumberg Flagle has discussed the decision processes involved in balancing demand and cost of available service in any given area This requires knowledge of linear programming and stochastic theory if it is to be used to its full advantage in solving hospital administration problems Several approaches may and probably should be used and physicians, administrators, operations analysts and many others must converge in the study of local hospital practices and solutions found in other similar institutions

---

*Bill at end of chapter

LEGISLATURE OF NEBRASKA

SEVENTY-SECOND SESSION

Legislative Bill 326

FINAL READING

Introduced by Fern Hubbard Orme, 20th District
Read first time January 25, 1961
Referred to committee on Public Health and Miscellaneous Subjects
Sent to printer January 26, 1961
Final form sent to printer April 17, 1961

A BILL

FOR AN ACT relating to public health and welfare, to provide for confidential information received by the State Board of Health, Nebraska State Medical Association, allied medical societies on in hospital staff committees in the course of a medical study for the purpose of reducing morbidity or mortality, to provide that such information and material so furnished may be used only for the purpose of advancing medical research and medical education, to provide for general publication of a summary of said studies, to provide an exemption from legal liability for those furnishing such information and for those studying and publishing the results and summaries of such studies, and to provide that such material and information and any findings or conclusions of such authorized groups shall be privileged communications which may not be used or offered or received in evidence in any legal proceeding

Be it enacted by the people of the State of Nebraska, Section 1 Any person, hospital, sanitorium, nursing or rest home or other organization may provide information, interviews, reports, statements, memoranda, or other data relating to the condition and treatment of any person to the State Board of Health, the Nebraska State Medical Association or any of its allied medical societies or any in-hospital staff committee, to be used in the course of any study for the purpose of reducing morbidity or mortality, and no liability of any kind or character for damages or other relief shall arise or be enforced against any person or organization by reason of having provided such information or material, or by reason of having released or published the findings and conclusions of such groups to advance medical research and medical education, or by reason of having released or published generally a summary of such studies

Sec 2 The State Board of Health, the Nebraska State Medical Association or any of its allied medical societies or any in hospital staff committee shall use or publish said material only for the purpose of advancing medical research or medical education in the interest of reducing morbidity or mortality, except that a summary of such studies may be released by any such group for general publication In all events the identity of any person whose condition or treatment has been studied shall be confidential and shall not be revealed under any circumstances

Sec 3 All information, interviews, reports, statements, memoranda, or other data furnished by reason of this act and any findings or conclusions resulting from such studies are declared to be privileged communications which may not be used or offered or received in evidence in any legal proceeding of any kind or character, and any attempt to use or offer any such information, interviews, reports, statements, memoranda or other data, findings or conclusions or any part thereof, unless waived by the interested parties, shall constitute prejudical error resulting in a mistrial in any such proceeding

Chapter III

DIAGNOSIS

WHEN physicians are first confronted with the idea of a computer in some aspect of clinical practice, they often focus their attention on the use in differential diagnosis This is an aspect of potential computer application about which physicians seem to develop either enthusiasm or strenuous opposition One opinion is that the vast majority of patients have problems that are so simple to diagnose as to obviate computer use and that those patients who have difficult diagnostic problems will demand careful observation by the experienced physician in order to obtain data of diagnostic importance and that the important information is from a type of extrasensory perception that can neither be defined or communicated to the inexperienced let alone to a cold machine

Some physicians are in the habit of consulting books listing signs and symptoms and related diseases when they have an especially difficult diagnostic problem Physicians attending clinical pathological conferences, especially if they must discuss the differential diagnosis and the logic of reaching an opinion as to the disease process and the complications most likely present in the patient, are well aware of the mass of literature and experience that is available for review, even when the problem has been narrowed down by the limited signs and symptoms of one patient There must be a significant number of patients whose diagnoses could be assisted by automatic data processing Diagnosis need not be limited to finding a single disease but should include a detailed appraisal of possible subsidiary conditions, weaknesses and potentials for disease Physicians need to know all they can learn about their patients and machines do help in gathering and arranging data for more efficient use Vannevar Bush, in 1953, observed that too much was required of memory and not enough time was left for synthesizing thought He suggested that electronic means be exploited to take over the data storage and retrieval areas in medicine

Nash, 1954, observed how inadequate a textbook is for making the many comparisons required in differential diagnosis He experimented with a special ruler with about 300 diagnoses listed on the side and multiple inserts each labeled as a sign or symptom and with lines marked opposite diseases when the sign or symptom was present in that disease The diagnostician would use such an aid by selecting those inserts corresponding to the observed positive signs and symptoms and when some disease was found with the lines across all or most of the inserts a diagnosis could then be considered likely

Such a comparator is illustrative and instructive It is a landmark in a shift of thinking from books on differential diagnosis to automatic data processing and storage equipment for the busy diagnostician, but all such methods are primarily aids and will still depend on the experienced physician whose senses are trained to collect the most pertinent data with great skill and some accuracy Physicians freed to observe and converse with their patients would feel that their time was used in a more productive way If they recognized a disease at a glance no source material would be needed When the diagnosis is difficult the accumulated knowledge should be more available and in a quantitative and more certain way

In 1957, Lipkin and Hardy reported a study of the data characteristics of hematologic diseases as coded for application to marginal punch cards The data included information from the case history, physical examination, and laboratory examinations Standard textbooks on hematology were consulted during the formulation of the problem and master cards were prepared characteristic of each hematologic disease Considerable preliminary planning and preparation was carried out with marginal McBee punch cards, and the results of these studies have been published More recently, Lipkin and fellow workers have transferred this problem to a digital computer

They use three magnetic tapes Tape A has the case history data from each patient, and the first step is to compare this material with Tape B which contains definitions of classes of disease Twenty diseases were chosen for study and these were grouped into nine different general classes The first class, for example, contains the diagnoses pernicious anemia and tropical sprue Tape C contains definitions of each disease within a class and after the comparison has been made between Tape A and Tape B, then the appropriately suspicious classes are selected from Tape C for further comparison with the patient and further differentiation of the diagnosis from class to specific disease results

The program is designed to make a positive diagnosis when all of the critical items of a disease, as defined in the Tape C, are satisfied by a given patient's data If some critical item has been tested and found absent, the disease is still retained in the differential diagnosis and the appropriate information is printed out suggesting that the physician recheck this item If two or more critical items are found to be absent, the disease is discarded, but the printout explains what information was absent If one or several critical items have not been tested, then the program lists information which is required before the probability of a diagnosis can be stated A weighting technique is used for the signs and symptoms and laboratory findings in the disease definitions and important negative findings are also weighted The construction of disease classes serves to reduce computer processing time by an estimated 50 per cent

In this particular study, the diseases are fairly well defined, but, if such a program were expanded to include other conditions, a problem might arise due to the vagueness in the definition of certain diseases As this program is given a more thorough clinical trial, it will be possible to alter the weighting factors by appropriate statistical techniques and to add additional diseases and classes of diseases as it becomes evident that these are required

Lipkin has indicated that it is desirable to list the diseases in the differential diagnosis in sequence of decreasing probability He also emphasizes the possible usefulness of a program of this sort as an aid in teaching differential diagnosis and that its use in medical practice might contribute to an increasingly analytical (less intuitive) basis for diagnosis If the computer were used to store large amounts of factual information which now a physician must gather from articles and textbooks or from his own memory for each diagnostic problem, then perhaps the physician's time could be concentrated more in the area of original thought and less in the area of a filing clerk

Warner, Toronto, Veasey, and Stephenson prepared a matrix of probabilities based on a study of 1,035 patients with suspected congenital heart disease who were referred to their laboratory for diagnostic catheterization and other special studies They listed *fifty symptoms and signs and thirty-three diagnoses* The incidence of each symptom in each disease was transferred to punch cards and stored in the computer along with the program for carrying out the analysis For each subsequent diagnostic problem patient analyzed, the symptoms from a check list were punched on a card which was run through the computer A conditional probability equation was then worked out automatically with the appropriate probabilities from the table and finally the diseases with probabilities greater than one per cent were printed out together with their respective probabilities

Thirty six cases were subjected to this program The "true" diagnosis was made by physiologic studies and observations or at the time of operation The material was also submitted to three experienced cardiologists The result was that the computer diagnosis was in agreement with that of the cardiologists The differential diagnosis from the equation was frequently more complete and appeared more logical in retrospect to the clinicians than the differential diagnosis listed by each of them prior to seeing the equation's prediction These workers now plan to use the computer to recalculate the probabilities periodically as additional data accumulates in the further use of the program In a sense, then, the program will "learn" from experience by virtue of refinements of the table based on the accumulating data It is also possible to study the effect of modifying the symptom set for a given patient These workers report an enthusiastic acceptance of use of this technique in the teaching of medical students

The formulation of a diagnostic problem in this framework requires a very exact definition of each symptom and each disease It requires an understanding of which symptoms are exclusive and which ones are interdependent The method requires that only one sign or symptom be used from several having the same information content The important determinants of the diagnosis are emphasized and symptoms which have little information content can be recognized and eliminated A student or physician testing his collections of data in such a study should develop a more acute sense of the relative value of the data that he collects and should become more efficient in his history-taking and physical examination He should learn to ferret out those items which have the greatest information content

So far the approaches discussed relative to differential diagnosis have followed a pattern of comparing diseases with signs and symptoms The refinements have consisted primarily of refinements in data handling and in the use of probability statistics These studies demonstrate what a tremendous amount of information must be collected and stored in tables even for a very narrow area of medical practice There is no question but that programs can be developed by this means if enough interested and highly trained physicians throughout the country will concentrate on defining diseases and their related signs and symptoms, and will collect the necessary data for preparation of the reference material both from the literature and from further prospective protocol studies of patients There is the problem of differences in geographic effects and the influence of epidemics A matrix might continually change, less rapidly for certain broad types of diseases such as congenital defects and more rapidly for infectious diseases It is common for physicians to make rather difficult diagnoses consistently and occasionally in error simply because there has been an outbreak of that particular disease in the area

There may be ways of developing differential diagnosis programs which do not require as much searching of unimportant data in order to focus on the most pertinent aspects of the problem Newall and Simon have suggested a means of study of human thinking which is somewhat analogous to simulation experiments which will be mentioned in a later section Clynes, for example, constructed a computer with parts representing the various components of the respiratory-heart rate reflex mechanism and placed man and machine side by side Recordings of respiratory movement were taken from the man, fed through the computer, and the computed heart rate was then compared with the patient's actual heart rate Adjustments were made in the machine until it corresponded with the human Whether or not the analog was a unique solution, the very construction of it and testing of it under a variety of conditions led to new ideas about the way the reflex system works This is mentioned here in some detail because it serves to introduce the suggestion of Newall and Simon They would make a somewhat similar comparison of problem solving by man and the problem solving analog called GPS (The General Problem-Solving Program) GPS has not specifically been designed for use in differential diagnosis It is a very general program It is a possible analog of the way humans solve problems It is software which can be used with any moderate sized digital computer There is no assumption of similarity between either the brain and the GPS program or between the brain and the computer used Newall and Simon have suggested that we observe how GPS and a human subject solve the same problem Both the brain and computer programed with GPS are devices for manipulation of symbols and for problem solving behavior To watch man and programed machine work on the same problem would have the virtue of developing a specific diagnosis program useful in medicine while, at the same time, exploring the way in which an experienced and capable physician functions in making his differential diagnoses

The drawback to this approach is that a rapid easy solution applicable for early demonstration of clinical use would be unlikely A great deal of study of physicians' logic and modification and amplification of the problem-solving program would probably be required The large electronic storehouses of medical knowledge would still be needed, but perhaps they could be used more efficiently Perhaps we could find out to what extent the intuition of physicians is simulated by a probability model Certainly physicians do not carry around in their heads large matrices of correlation coefficients between all possible diseases and all known signs and symptoms Such a storehouse is, in fact, even beyond the capacity of any computer yet designed

The human brain seems to be capable of developing very clever ways of figuring out the relationships needed from a very limited amount of actual memorized material Physicians think of diseases in a dynamic context as a sequence of changes occurring in certain organs, and they remember the signs and symptoms because of their understanding of the mechanisms which bring about these *indices of the hidden disease process They do not usually memorize* long lists of signs or symptoms any more than we remember stories by memorizing a list of nouns, verbs, adverbs, etc There are many ways of classifying diseases on the basis of anatomy, physiology, bacteriology, pathology, etc The clever diagnostician remembers many little relationships because of some common denominator which may be rather obscure on superficial examination of the actual logic used at the bedside This is the experience *of which physicians are proud This is what appears to be artistic* ability in the practice of medicine and why it may seem to many physicians that they are gifted in making correct decisions without definable data and why they profess that rare diagnoses can be correctly made when no statistician would dare stick his neck out on the basis of the data presented

There is, then, a large job to be done by the student of physician logic or of human brain problem-solving We must find out how the best diagnosticians *function*, what *information* they use and how they use it to make the astute diagnoses that they make There must be also, at the same time, review of the way in which physicians make mistakes The preparation of programs by this method will be carried out in occasional spurts of brilliance but *chiefly through long, hard, seemingly unrewarding labor* Contributions will be made by individuals and groups of many interests and levels of training It will require considerable teamwork, but the resulting program will contain the fruits of all of this labor and better diagnoses should ultimately be possible with the programed cumulative brilliance of all of the best available brain power How soon a workable medical GPS will be available depends on our interest and effort.

Lusted has estimated on the basis of some preliminary work done in hematology that the necessary work required to examine the entire body of medical literature and to collect new data on the basis of carefully defined standards would be in the range of

500 to 600 man-years The rewards of such effort might be, according to Lusted, the opportunity of a physician to contact his regional center, supply the available data, and request an immediate differential diagnosis list of the most probable diseases together with information as to what additional examinations might be necessary to make a definite diagnosis He would also be able to obtain the latest information on the treatment of the disease and its expected prognosis Lusted suggests that the final mathematical model of medical diagnosis may involve a sequence of a number of different approaches, including (1) logical analysis, (2) probabilistic analysis, and (3) value theory to assist in the choice of a treatment plan

It appears at the present time that the student of computer diagnostics must have a background in probability theory and that he might be helped by the logical algebra of George Boole or, according to Swisher and Agalides, by the related symbolism of McCulloch-Venn and Karl Menger's Functor algebra The symbolism finally used by any one individual worker depends as much upon the influences of past experience and education in logic as it does on the diagnostic problem at hand. It may also be influenced by the worker's acquaintance with computer techniques A computer technologist working with analog computers might prefer a symbolism which was more immediately transferable into diagrams of switching theory The digital computer expert might be more inclined to use conditional probability, multiple regression or factor analysis techniques The statistical approach of multiple regression would be helpful in setting up the *a priori* probabilities given a sufficiently large number of case studies and would then be applicable to the processing of an individual patient's diagnostic problem through the use of these probabilities The present state of the art indicates that computer diagnosis programs are practical and could be useful tools for the busy physician or for the physician confronted with difficult diagnostic decisions

Before such programs become an actuality for many of the disease categories and for specific geographic regions there must be a tremendous amount of work both in development of programs and in the collection of large volumes of reference information regarding actual patients with known diseases of all kinds For certain data such as routine laboratory tests, the recording of data offers no problem There are problems in the collection of data from history and physical findings in a form which can be conveniently stored and at the same time will be acceptable to the physician who must make the initial observation and recording There are some observations which a machine can make more accurately than can a physician

There are also some observations that a physician cannot make which require a computer such as the electrocardiograph To the extent that the physical examination is automated, decisions will need to be made as to the kinds and amounts of data which are to be recorded and a good part of the computer application will probably be on line and dynamically programed so as to retain only observations of diagnostic significance and to avoid the storage of large amounts of normal data which might conceivably contain discriminatory information but which the busy physician would never get around to analyzing

Whatever approach is used physicians will be made more rather than less necessary since they will need to make the important decisions, collect data that no machine can collect and contribute to the program development that determines the usefulness of computers The efficiency of any program will be related to the intensity with which it is used in the local situation since many ill defined ethnic, social, psychologic and epidemiologic factors may influence the pattern in which a standard set of questions are answered The conclusions reached and programed into an automatic interpretation of the Cornell Medical Index or the Minnesota Multiphasic Personality Inventory, for example, must be reviewed in many different circumstances

This is one of the areas of clinical research which has been shown to be feasible but which will need implementation and co-ordination in numerous centers and probably in satellites of these centers Sets of questions have been formulated Simple acceptable methods of having these questions answered have been worked out Programs have been written for automatically processing the tests and evaluating the results Validation of results remains and will always remain an important and exciting task for physicians The physician must decide upon the final criteria in any given instance as to the accuracy of all analyses whether they are carried out by physicians with their unaided senses or through intermediate machines and laboratory workers The computer is one more tool to help in the never ending struggle to know and to control our environment and the diseases that afflict us

Brodman, van Woerkom, Erdmann and Goldstein reported in 1959 on a study with the Cornell Medical Index (CMI) which would seem well worth while repeating in any center where physicians are interested in contributing to and becoming involved in the automation evolution that is beginning They collected from nearly 6000 consecutive male patients as an initial sample Some sixty diseases were then studied and the pattern or score on the CMI was computed for each disease Data were then collected from 2,745 consecutive patients beginning January 1, 1956 and on the basis of the previously computed probabilities of each complaint for each disease the new patients were assigned likelihood indices for each of the sixty diseases studied A diagnosis was considered as made when the index reached a certain level The computer diagnosed 44 per cent of patients having the sixty diseases and identified the organ system diseases in 55 per cent When a random sample of 350 questionnaires were examined by both computer program and physician there was no statistically significant difference in the results except that physicians tended to diagnose psychoneurosis more accurately than did the machine

Swenson and others in 1962 have reported on their initial experience with a personality inventory test which could be used alone or in combination with the Cornell Medical Index and might bolster the weakness of the latter in determining psychoneurotic tendencies as well as pointing out major neuropsychiatric problems This copyrighted text which was developed by Hathaway and McKinley at the University of Minnesota and called the Minnesota Multiphasic Personality Inventory (MMPI) consists of 550 yes-no questions which were originally standardized on a group of normal persons visiting patients at the University of Minnesota Hospitals, persons in the employment of the Works Progress Administration, high school graduates who had not yet entered college and patients in the University of Minnesota Hospitals who were under treatment for nonpsychiatric illnesses

Patients were then selected who demonstrated rather clear-cut psychiatric problems which could be classified in one of eleven different groups, namely hypochondriasis, depression, hysteria, psychopathic deviate, masculinity femininity, paranoia, psychasthenia, schizophrenia, hypomania and social introversion Grading the test consists of determining whether an individual patient answers the questions in a way consistent with the original normal group or with one of the psychiatric group categories There are some additional keys which have been developed to validate the "test taking attitude" of the individual and are said to be reasonably effective in detecting low reading ability, malingering, lying, confusion or defensiveness

This test has for some time been machine scored at the University of Minnesota but the Mayo Clinic Study has added to this a program which automatically converts the score into those standard interpretative statements that are used by psychologists The following is an actual example "Tends to minimize or smooth over faults in self, family and circumstance Somewhat tense and restless Considerable number of physical complaints Prominent concern with bodily functions Moderately depressed, worrying and pessimistic Conscientious, orderly and self-critical Conforms reasonably to social codes and morals Has a combination of practical and theoretical interests Respects opinions of others without undue sensitivity Normal female interest pattern for work, hobbies, etc Has capacity to maintain adequate social relationships"

It is apparent that the latter portion of this example is describing a normal individual, but the initial statements, indicating abnormal patterns, could be helpful in preparing the way for a more rapid understanding between the patient and his physician The patients look upon the test as another effective means of communication which fills in the time that otherwise might be "wasted" while they are waiting for appointments with their physician or consultants or for other special laboratory tests and examinations The test is presented to the patient in the form of a small pad of cards and a special electrographic pencil and takes one to one-and-a-half hours to complete The cards are collected together twice daily and processed and the reports are printed out on a chain printer on paper which can then be inserted in the patient's record Each record requires about thirty-five seconds for processing

This approach of allowing the patient to take some of his own history by the use of standardized questions and mark sense cards has been amplified in the Tulane study to include forms which are made out by the physician, nurses, and technicians so that a great deal of the data which is usually found in the patient's hospital chart can be machine processed, assembled and in a limited way analyzed There is no reason why many of the physical findings usually written in a physician's record could not equally well be recorded by the use of a check list Information which is so specialized and unusual as to be unavailable in any standardized set of questions or check list could still be wirtten out and maintained in that form or transferred to the machine record by techniques which are quite similar to the use of an ordinary typewriter Any data that can be coded for machine processing and storage becomes available then for research in the field of diagnostics, the study of probabilities, factor analysis and stochastic processes

In such an arrangement each individual patient contributes to the information content of the whole organization This has always been true to the extent that individual patients contribute to the experience of the physicians caring for them, but this experience could become much more rigorous if added to accessible electronic storage and would no longer depend upon the fallible human memory *The physician in the future will have available such* extensive and accurate information that he can devote more of his attention to nonroutine decisions which no machine can accomplish

In the beginning there is no obvious reason why a group of physicians in a small hospital without computer facilities could not participate in such a study as a "satellite" with very little local investment in equipment This would have the added advantage for the central study of adding to the size of their sample as well as broadening the base from which the sample is obtained

Automation may make possible the earlier recognition of some medical problems through the use of mass screening Smith, Miller, Marienfeld, Hahneman and Willard reported in 1959 a study of

3,272 tape recordings of heart sounds in children in which a technician recorded approximately 250 children per day and a physician listener subsequently screened these at the rate of 140 per hour Seven children were found with definite heart disease and only two of these were previously known Gerbarg and others have recently shown that a digital computer can be programmed to analyze phonocardiograms but they feel that more work needs to be done in recording techniques in order to make full use of a full-frequency spectrum and to minimize noise Wherever the job done is repetitive and where fatigue and emotion might have some effect on human effectiveness, machine applications can be considered The machine is able to remember, compare, compute and make decisions on the basis of predetermined criteria

Chapter IV

ELECTROENCEPHALOGRAPHY AND ELECTROCARDIOGRAPHY

COMPUTERS are forcing man to see more of the art in science, more of the forest along with the trees, the patterns within patterns, the differences in viewpoints and interpretation and the difficulties of being really sure about any interpretation Current *investigation of electroencephalographic tracings, electrocardiograms*, photoroentgenograms and other patterns will be discussed in an effort to emphasize the need for a pluralistic approach in analysis and the benefits to be derived from computer use in these endeavors

Most of us for years have had an oscilloscope at home with which we allow our children to play (TV) and each of us has a portable 10 watt computer that is priceless, infinitely complex and not too accurate in our respective cranial vaults Man carries with him *in his five senses the peripheral transducers and the network* for obtaining data plus all of the servomechanisms required to keep the system working at a level that may provide useful maintenance, reproduction and occasionally some creative output

Once a year a few thousand of these complex living systems gather together in a large broadcasting studio in Des Moines, Iowa for an All-State Music Festival to listen to an orchestra of 200, a band of 270 and a chorus of approximately 600 high school musicians It is interesting to sit in one of the seats in a far corner of the balcony and think about the complex data flow, analysis and "creation" that is available in such a spectacle and what a wealth of material the investigator in electroencephalography has at his command No individual in the entire amphitheater receives exactly the same information None of the messages are pure The hall acoustics, though good, are far from perfect There are some baffles hanging from the ceiling which probably reduce some of the reverberations

Sometimes the 1000 performers and the audience seem not unlike a large model of a single brain with a diffuse alpha rhythm and superimposed more melodious and informative strains winding their way through the horns and the strings to the finale which is climaxed by a raucous discharge of clapping from the entire "lower" centers This analogy is also not unlike a huge computer with the conductor acting as the control and the music on the stands and programs in the hands of the audience functioning as the software or program for this particular run But there is more to the analogy The observer in the balcony is aware of intruding television cameras and microphones which probe and record the entire operation and then less than 24 hours later we sit before our oscilloscopes, each in his own laboratory, and attempt to obtain some of the original experience For the observer of the oscilloscope who was also present during the original performance, it is obvious that a great deal of distortion has crept into the experiment There is a waxing and waning of sound The striking colors of band uniforms mixed together from the entire state and the colorful formal dresses are now black and white and the clear detail is lost except on the close ups where it must be confessed that the oscilloscope gives considerably more detail than could be obtained from a corner balcony seat Some of the loss of the original signal developed in the recording and some in the rebroadcast We can identify some of the artifact as due to the periodic running of an electric train in a room above the television set or the over flight of an airplane There also seems to be a change which depends upon the characteristics of the original program The orchestra seems to have been reproduced the poorest, the choral group the best and the band somewhere in between Perhaps the choral group has the best chance of producing a very clear and concise signal and a collection of stringed instruments is more difficult to control

All of the problems mentioned in connection with the reproduction of a musical event exist in the use of electronic equipment to record and analyze the electrical activity of the brain The electroencephalographic tracing is a collection of the loudest signals circulating in some 10 billion cellular elements and the propagation of each individual signal involves its successive hundreds of neurons Every sensation of which the individual is aware from his periphery and his internal transducers has its counterpart in the brain There are, in addition, the facilitated spreads of activity through neuronal patterns of memory The electrodes recording from points on the scalp are separated from this wealth of data by a filter of bone and soft tissue The electronic equipment components each add their own impedance and distortion

It is not surprising in view of the complexity of the problem that limited progress was made in analysis of electrical activity of the brain until statistical methods and computer technology were recently applied to the problem Human endeavor and brain function are primarily statistical in character It has been said that God (and we might add mathematicians) study the universe and have no need for statistics (or computers) Ordinary men study only samples of the universe and their creative energy is spent in a continuing research of old ground in an effort to understand phenomena as if they were observable in their true mathematical perspective Ordinary men need all of the help that they can obtain from computers, statisticians (and mathematicians)

The research worker is continually searching for some unexpected rhythm in his recordings and in the process he, of necessity, must lose most of the real information containing data which he calls "noise," variance or "hum" When the investigator finds some unexpected rhythm he immediately begins to analyze it to see how much information is really hidden therein, but because it is a rhythm and to the extent that it is repetitive, monotonous and redundant, it ceases to have any additional information content It is a paradox that man strives continually toward recognition of order in the laws of nature and yet information theory relies upon discovery of the unexpected The unexpected is, however, a new predicability in rhythm and not more noise, chaos or entropy

Many theoretical models are used in an effort to explain the basic functional design of the brain as observed in the experiments which computers must help process In addition to the experimental models there are other models which we call statistical methods These are mathematical models and each requires certain well defined characteristics in the distribution of the data in order that the derived statistics be valid, unbiased estimators Whether a data source is sending information in a normal, log normal, Poisson or some other distribution, if the statistical model specifies a normal distribution of data a transformation may be in order before the statistical method is applied An average or *standard deviation is uninterpretable if the data has a marked skewness*

The desirable transformation should probably be derived from data collected from the brain during appropriate control periods of the experiment and this is one of the points at which judgment must be used by the investigator The entire application of statistical methods implies repeated judgments of the research worker There are some statistical models that are not dependent upon the type of distribution and even a skewed distribution can be analyzed with Gaussian statistics if the spread is not too great Whenever the standard deviation is as large as the mean there must be some question about the use of non normal, untransformed data Whenever possible the investigator should retain the original data or at least a portion thereof for study and display along with some of the derivatives and conclusions so that the extent and some *of the nature of the loss of information can be observed*

One test of the specific analysis which is programed or wired *into a computer is the accuracy with which derived statistics can* be used to regenerate an original signal (without the original noise and with a minimum of new noise) This can never be completely accomplished because of the inevitable loss of information that comes about in every reduction and transformation of the data This type of check of the hypothesis regarding specific procedures of data reduction is somewhat analogous to the use of Koch's postulates in a bacteriologic proof In bacteriology the isolated infectious agent must be carried through cultures outside of the host and then the disease must be reproduced by the isolated agent before it can be accepted as the basic causal factor in the original disease

In the brain messages are received from many types of peripheral transducers sensitive to different modalities of heat, pressure, vibration, pain, sound and light The signals undergo a programed analytical breakdown into essential and significant components and there is then resynthesis into a complex projected multisensorial image which seems to be at a given position in the four dimensional time-space that man occupies Computers are programed to function somewhat like their human counterpart, to take complex signals apart and to resynthesize them Machine computers are much less redundant than the brain because they are composed of finite pathways and function in a nonstatistical manner, even though many of the problems computed are statistical A message through a machine follows along single wires and through isolated elements There are a few redundancies in the form of checks of the computer equipment and checks for some of the most common mistakes that come about in human programing The machine computer, however, has nowhere near the redundancy of the brain computer A message coming through the brain is not unlike the onrush of traffic at a busy intersection when the light changes to green It involves successive hundreds of neurons and each neuron is connected to a hundred more in a network of potential redundancy There are thresholds, counteracting neurons and enzymatically controlled chemical intermediaries between the neuronal synapses which help stabilize human behavior and make possible continued function even at times when relatively large segments of the neuronal population are completely destroyed To really progress rapidly in study of brain function or cybernetics, as Norbert Wiener has termed the study of control in man and machines, there must be multidisciplinary teamwork and pluralistic methodology

Because of the statistical problems involved computerized electroencephalographic study will be of continued interest throughout the biomedical and behavioral science Whereas the biochemist, depending upon the nature of his hypothesis, may have to work for months to obtain enough data to justify one analysis, the electroencephalographer is literally drowned in data with the flick of a switch It would be safe to predict that the electroencephalographer will become the applied statistician supreme It is not surprising, therefore, that the electroencephalographer, in his unguarded moments, refers to the problems of the electrocardiologist as being trivial Much of the problem up until now for the electroencephalographer has been to obtain a recognizable signal, a predictable pattern within the noisy onrush of electrical activity which originates throughout the nervous system but focuses in the brain

Whereas other research workers have been able to use the large and somewhat inaccessible computers of business, the brain research worker has required and built special equipment for on line computer applications in the laboratory One common technique is to use a stimulus to the experimental subject which also triggers the control section of the computer and from selected electrodes on the scalp the recording is shown on an oscilloscope With one stimulus there is no identifiable effect in the tracing, but then a second identical stimulus is given and the oscilloscope shows the average of the two tracings As the stimulus is repeated over and over again, the averaging of the curvilinear electroencephalographic tracing becomes less and less noisy and more and more a definite signal which is related to the stimulus at zero time All activity from the brain which is unrelated to the stimulus is considered as noise to be averaged away and the operator continues the experiment until the signal/noise ratio reaches the desired level

A second type of computer is designed to compare tracings from two points on the hemisphere or from both hemispheres in a subject or from two subjects who are receiving the same stimulus This computer has a mechanism for delaying one of the signals, multiplying the two signals together and integrating the result from a selected segment of the tracings The recorded brain waves are played over and over with varying time delay of one of the tracings in order to find the appropriate time delay which will give the maximum correlation between the two tracings One of the objects of such a study is to be able to trace the propagation of a signal between different recording sites

Once signals are separated from the background of unwanted electrical activity it is common then to compute the amplitude at varying time intervals following a stimulus and to study the relationship between various amplitudes and the latency One of the problems of averaging which is also used in correlation studies is that the subject is responding repeatedly to the same stimulus which means that the experimenter is investigating a rather monotonous form of brain function Some of this problem is perhaps eliminated by using a randomly repositioned light spot on an oscilloscope and recording some additional information from the periphery as to the position of the individual's eyes, the orientation of the head in space as well as the position of the object being tracked by the experimental subject Barlow, using this approach, averaging techniques and scalp electrodes, observed that it was possible to separate the electrical activity of the brain related to shift in spot position from that related to changes in eye position

The amplitude of the electroencephalographic activity is larger during sleep as can be seen in the original tracings Study of the pattern of rhythmic burst activity, or the underlying alpha rhythm, is less obvious in original tracings and is helped by computer techniques It has been possible to show that when an individual is placed in a dark, anechoic room that there is a progressive decrease in the abundance of rhythmic burst activity which is detectable during as brief a period as ten minutes It is also possible to drive the rhythm of the brain with a flashing light so that two subjects can become locked into the same phase It may be that this phase activity is a mechanism that the brain has for erasing some of the noise and preparing the brain for association functions required in higher thought processes If all activity is eliminated by some means no thinking can take place The brain is again not unlike electronic or inertial equipment which is easier to regulate or modulate in a state of motion than from a state of suspended animation

It is apparent that as more is learned there will be applications of computer analysis of electroencephalographic data in clinical applications in diagnosis The electroencephalograph has already been used to trigger a servomechanism for the regulation of the administration of anesthetic agents It is to be expected that a series of statistically derived profiles for help in diagnosis of various psychiatric disorders will be greatly facilitated by computer data processing The computer then, in summary of electroencephalographic work, has the double duty of assisting in identification of the signal in a background of noise and then once the signal has been identified a second analysis is required to determine the amount of information available in the tracing There are special studies being made of neuroelectrical activity in such organs as the olfactory bulb and its correlation with central activity of the brain

A special field of investigation is built up around electroretinography, recording and analyzing data obtained from the surface of the eye Certain parts of the observed pattern seem to be produced by the cones and other parts by the rods It has been possible to find useful differences in the patterns produced in patients with varying degrees of retinitis pigmentosa by using a light at a position in front of the patient and then taking recordings from both eyes with the patient looking at fixed angles to the right and left These latter applications were made by Peter Guross and use the principle of evoked response for evaluation of visual function in a clinical problem

Harold Shipton at the University of Iowa has built a computer which belongs to the family called toposcopes This instrument records simultaneously from sixteen or more electrodes attached to different areas over the surface of the scalp, automatically computes correlograms from each combination of electrodes and displays these on sixteen cathode ray tubes so that time is represented by the continually moving clockwise radians and the point of correlation in time can be related between points on the skull through visual inspection or recorded photographically This is a technique used in submarine search and the instrument has some characteristics for information recovery that exceed the predicted and that have not been explained This type of instrument exploits the complexity and volume of signal available in the brain and adds a broader geographic or topologic dimension to correlation in time

There is another type of computer use that will be mentioned in greater detail in the next chapter where systems, feedback and transponder functions are reviewed but which deserves special attention in relation to electroencephalography The brain is the center of control for many body systems, for internal homeostasis and for proper orientation of man in his external environment An important aspect of this is feedback and Shipton, working with F L Golla and W Grey Walter and others at the Burden Neurological Institute, has also been an early contributor to this area through design of "trigger" circuits which allow the experimenter to introduce positive feedback and to throw the brain into oscillating activity which can be used for identifying patients with "epileptic tendencies" as well as for investigative studies

This "trigger" circuitry functions to receive, smooth, amplify and transmit basic patterns of the brain waves so that a stimulation signal such as a flashing light can be fed back to the subject in any desired time delay relationship to recorded ongoing brain rhythm Normal subjects or patients with only minor psychiatric disturbance report that such a feedback flashing light causes varying sensations Some report a subjective sensation of change in color of the light Others note a sensation of movement or "swimming" Mood of the patient influences the type of response of a given individual It has been suggested by Golla, according to Shipton, that long exposure to such stimulation might succeed in effecting a permanent change in the mental state of a subject

If electroencephalographic analysis is analogous to attempting to isolate an instrument in an orchestra while watching a delayed television broadcast, then electrocardiographic analysis is analogous to listening to a high fidelity stereophonic tape recording of Heifitz The first human electrocardiogram was recorded by a London physiologist, Augustus D Waller, who introduced the term "electrocardiogram" into science The first clinical investigation was carried out by Thomas Lewis in England An important early paper *was published in 1913 by Einthoven, Fahr and de Waart* and was titled "On the Direction and Manifest Size of the Variations of Potential in the Human Heart and on the Influence of the Position of the Heart on the Form of the Electrocardiogram"

The heart is a muscular organ situated in the center of the chest It is a paired organ, the right side receiving blood from the body and projecting it into the lungs, while the left side receives blood from the lungs and propels it out into the systemic or general body circulation The upper two chambers are called atria and receive the blood The blood then passes into the more muscular ventricular chambers There is a septum between the two sides of the heart and the site of origin in electrical depolarization which stimulates the rhythmic beat of the heart is located in this midline septum There is a sino-auricular node (S A N) where the beat normally begins for propagation over the auricular walls and this stimulates contraction and propulsion of the blood into the ventricles The impulse then begins at the auriculoventricular node (A V N) and passes down the septum between the two ventricles in the special conductive *bundle of His* and thence out over the walls of the ventricles to cause contraction of the ventricles

The heart rests on the dome of the diaphragm in such a way that with deep diaphragmatic breathing the heart is caused to rotate within the chest The electrical impulse probably passes through the heart in what can be thought of as an eccentric double layer of negative charges on the inner surface of the chambers and positive charges on the outer surfaces with depolarization representing the propagation of the impulse The impulse is propagated from the heart out through the body to the periphery where it can be recorded with attached wire leads and a string galvanometer

Einthoven thought of the signal as approximating what would arise from a single dipole located in the region and position of the heart or a heart vector In 1912, George Fahr recorded heart sounds simultaneously with electrocardiograms and these and other studies revealed the relationship of the P-wave to contraction of the auricles and of the QRS and T-waves as related to the depolarization and repolarization of the ventricle

It was apparent to these investigators that the position of the heart influenced the form of the electrocardiogram, that heart rate was probably related to vagal tone, and that there was an associated variation in the pathway of excitation with changes in rate They observed that hypertrophy of the cardiac muscle caused changes that were similar in many respects to the changes caused by shift in position of the heart One of the most common abnormalities of aging man is coronary artery insufficiency and occlusion of various branches of the arteries with consequent impairment of conduction Sometimes the infarction involves a branch of the *bundle of His* and so called bundle branch block occurs Occasionally complete heart block is present Also there may be abnormal placement of the site of origin of the impulse in the auricle There is also a variety of disturbances in rhythm

It has been customary to attach electrodes to the extremities and to use an additional electrode placed at various positions over the chest for some of the readings and to record varying combinations of these electrodes for a total of some twelve standard leads The voltage is recorded against time on paper The analysis is made by measuring the heights of various components and the widths and distances between these components as well as studying the shape of the waves For certain diagnoses some of the leads are much more valuable than others The reader has some information about drugs that the patient is receiving, especially digitalis preparations or other drugs which might affect the electrocardiogram and also usually has the blood pressure recorded and a little information as to the type of problem which the clinician ordering the electrocardiogram is looking for The cardiologist also is familiar with the common artifacts which must be disregarded as noise such as 60 cycle current and muscle tremor Although some measurements are made, these are rather inexact and most of the interpretation is qualitative rather than quantitative

To all of the other variables which affect the interpretation there are the added variables of the complex human electrocardiologist such as training, experience in reading the tracings, alertness or fatigue, and emotional bias Validation or feedback is haphazard so that the cumulative accuracy of the opinions is impressionistic A good electrocardiologist does have a vast amount of information from carefully controlled studies in which the deviations from normal have been studied in select groups with autopsy confirmed abnormalities

There have been careful studies made with models such as the one of Brody and Romans where electrically conducting material in the form of sheets of paper were cut out in the shape of a human body with a dipole constructed in the position of the heart and additional superimposed sheets of paper representing the heart tissue, lungs, abdominal viscera, etc There have been studies with recordings from within the esophagus and within the heart itself and within a bronchus near the heart

Nevertheless, when studies have been made of observer interpretation of electrocardiograms, there is good agreement among the cardiologists only on the most marked abnormalities There is surprisingly wide difference of opinion in interpretation of some of the tracings Epstein studied the diagnoses of five different observers using 451 records from young Spanish soldiers and found that the individual observers even made between two and nine reclassifications of their own original readings These variations were due more to interpretations than they were to objective description of the criteria for diagnosis

Segall studied the reports of twenty physicians using a set of 100 electrocardiograms and observed "the degree of variation in opinion is greater than I had anticipated in any group of men who work together and attend conferences together" There were twenty patients in the group known to have myocardial infarction, but one of the readers diagnosed myocardial infarction twenty-one times and another forty-nine times There was a majority of agreement in 77 per cent of readings

In a similar study by Davies with nine readers similar results were observed with complete agreement in only one-third of cases On the average Davies' readers disagreed with one out of eight of their original reports Davies concluded that the overlap of normal and abnormal tracings is much greater than is generally realized

Acheson observed that the only diagnosis which was uniform was that of atrial fibrillation There appeared to be a continuum between normal and abnormal electrocardiograms and between the various different diagnoses made of disease Acheson felt that perhaps the safest opinion to render was simply that the electrocardiogram showed coronary heart disease rather than trying to separate left ventricular hypertrophy, myocardial ischmia or myocardial infarction

In contrast to this, carefully programed computer analysis is *devoid of the day-to-day variation which is inherent in the spot type* of human endeavor The programs are written and rewritten by experts in electrocardiology, mathematics, statistics and computer technology, each individual adding his experience and each defect being corrected as it becomes apparent in the results of the analyses By virtue of these corrections the program is heuristic Another reason for the increased accuracy of the computer reading is that many more measurements of amplitude and latency are made together with the complicated mathematical analyses which would be prohibitively time-consuming and also liable to additional error if done by hand The interpretation can be printed out in any desired detail with some of the common measurements customarily reported The results of the analyses can be stored in computer memory along with information derived at a later time about the final diagnoses and outcome so that probabilities can be derived for each diagnosis The conclusions are qualitative and are, therefor, ideal for incorporation into larger probability equations containing data from other sources Pipberger has stated "From results *evaluated so far, it appears that the diagnostic accuracy of the* described computer methods considerably exceeds that of conventional procedures used in clinical practice "

Clinicians who feel concerned about the intrusion of computers into the analysis of electrocardiograms should be reminded that the electrocardiographic machine is already an analog computer and that data reduction begins in the sampling and lead selection

Since the output is in analog form and the analyses are carried out with stored program digital computers, one of the first requirements is analog to digital conversion The continuous line is scanned at selected intervals, the potential difference measured and converted to a number which can be recorded, usually on magnetic tape, but in a digital code The computer is then programed to identify some standard part of the tracing to provide consistency in the analyses

Differentiation is used The change in successive voltages is computed and this derivative, the time rate of change of the signal, is used to determine the point of the curve where the greatest negative rate of change takes place, or the R-wave The program *identifies other waves in relation to this point* Up until this time abnormalities of rhythm have been excluded from analysis and *separate programs are being written It is necessary, therefore,* to either prescreen the tracings or else to have instructions incorporated in the program to identify tracings with abnormalities in rhythm *When programs become available for analyses of disturbed* rhythm the two types of programs can then be placed in the computer and the program designed to automatically shift into the appropriate analysis, whether it be for regular rhythms or abnormal rhythms The power output is determined by integrating the curves ,

Standard analytical technique has been programed by Caceres, Steinberg and others as well as by Gustafson, Daniel, White and Staples (for children) Initial definitions of normal wave forms according to Caceres were arbitrary, but as precise as could be made from the clinical criteria usually used As experience is gained with the analyses these definitions have been revised and made more precise Sixty-cycle noise and skin and muscle electrical activity are eliminated by smoothing with a fitted nine point least square parabola The measurements of the identified wave forms are compared with the standard values and the logic in the program uses the results of these comparisons to instigate appropriate printing of statements as to the interpretation Statistical tables of variance are used to determine probabilities which are also incorporated in the diagnoses No analytical techniques are used that are unfamiliar to the physician who is used to reading these tracings except for the statistical addition Even with this standard approach Caceres reports that the reproducibility and precision of measurement are much greater with machine analysis

Gustafson observes that computer analysis is even of greater value in working with children in view of the changes that occur coincident with growth and development and the effect that this has on the criteria for diagnosing disease

Pipberger has observed that if only frequencies up to 200 counts per second were used it would be necessary to take 12,000 voltage readings for each second of recording if twelve standard leads are used There has long been a feeling that the multiple standard leads contained much redundant information and that if the heart is analogous to a dipole in three dimensional space that it should be possible to get all of the available information by recording three leads, each perpendicular to the other two

Scher, Young and Meredith studied this problem with factor analysis This is a statistical approach based in a matrix of simple correlations designed to determine the number of unique parameters in a large number of variables They recorded between eight and thirty-six simultaneous leads in seventeen normal persons.

The conclusion was that there were no significant voltages at the body surface in normal individuals which could be ascribed to more than three internal generators This does not prove that the heart can be considered as a dipole or heart vector, but it does suggest that all of the important information can be obtained by use of three orthogonal leads

Frank, and Schmitt have both described ways of obtaining these leads Frank uses one lead near the back of the neck, a standard left leg electrode and a series of leads around the chest at about the fifth interspace Two of these are lateral Two are anterior with one of these making an angle of 45 degrees with the left lateral lead The one posterior lead is on a line from the right anterior lead through the center of the chest

By appropriate placement of resistors in the wiring and combinations of these leads, a recording of three orthogonal or unrelated vectors is accomplished and Frank feels that these are optimum for the necessary compromises that must be made between a sound theoretic basis, accuracy, reproducibility, signal-to-noise ratio and speed of application The lead placement corrects for differences in torso shape and avoids the variation in left arms between individuals It is relatively insensitive to variation in the location of the ventricle due to variation in the diaphragm position or other causes

This was the climax of some five years of intensive theoretical and experimental investigation and was an attempt to correct the many known errors in existing methods of recording which seemed necessary before advances could be made in the quantitative area of analysis The great advances, both in the laboratory and clinically, will come from wide application of vector geometry and matrix algebra and various techniques referred to in mathematical name dropping as vector analysis, Fourier, harmonic and tensor analysis, use of orthonormal exponentials, etc

A vector is like an arrow, it has both direction and length The propagation of the electrical activity in the heart can be thought of as an arrow originating at one point and directed normally anteriorally down and toward the left By the use of the three unrelated (right angle or orthogonal) leads it is possible to measure the QRS vector and to use the Pythagorean theorem which says the square of the hypotenuse is equal to the sum of the square of the sides If then the squares of the three vectors are added together and the square root obtained the result is a measurement of the spatial magnitude of the QRS vector If a similar measurement is made for the T-wave it is then possible to calculate the angle between the QRS and T-waves

If the analytical technique is carried further and so-called polar vectors and eigenvectors are used it is possible to represent the propagation of the electrical impulse in the region of the heart, not as an arrow but as a vector loop which moves around in a plane at right angles to the polar vector If looked upon from the side the vector loop is shown to vary above and below the plane to give the appearance of a figure eight The polar vector identifies the orientation of the plane in space and the eigenvectors and their ratios indicate the configuration of the loop Vector analysis allows study of the directional and scaler (or power) components separately and combined

Pipberger studied fifty-six patients who had histories of old myocardial infarction and found abnormalities in the power (or scaler) analysis in 43 per cent of the cases Abnormalities in direction (or vector) components were observed in 95 per cent of the cases and when both amplitude and direction were used abnormalities were found in 98 per cent of the cases It was possible to study separately the QRS complex and T-waves as well as the combination of these components which is called the spatial ventricular gradiant

Pipberger carried out another study comparing the standard twelve lead electrocardiogram with the three, corrected, orthogonal lead tracings using analyses by physicians and found that there was a difference of 7 3 per cent of cases in favor of the twelve leads as far as recovery of information was concerned However, this was related to the clinician's ability to extract the information and was not from information loss in the three orthogonal leads When the recordings of the three orthogonal leads were used to resolve twelve leads similar to what clinicians usually look at there was then no difference whether the clinicians studied the original tracings or the tracings resolved from three orthogonal leads

Even computers have limitations in their storage capacity and there is a continual effort to avoid redundancy, to determine how much information is really available in any original signal and to obtain this information in a manner which will be useful to the clinician Spatial vector analysis would seem more acceptable once it becomes generally available than the twelve lead tracing In the case of left ventricular hypertrophy, for example, Pipberger has found that the progressive cardiac enlargement is associated with rotation of the polar vector in an anterior and inferior direction He has also studied 134 cases of myocardial infarction one or more years after the condition was first diagnosed in a patient After the twelve month interim a twelve lead electrocardiogram permitted diagnosis in 70 per cent of the cases In eight per cent of the cases the electrocardiographic readings were not definite enough to give a diagnosis and in 22 per cent of the cases the electrocardiogram indicated no abnormalities in the patients studied These same cases submitted to diagnosis by the more elaborate analyses were diagnosed 100 per cent as having had myocardial infarction Because orthogonal leads have a much smaller normal range than conventional leads, there is a better separation between normal and abnormal and this leads to greater accuracy in detecting the abnormal heart conditions Pipberger has analyzed over 3000 electrocardiograms using several mathematical curve analysis methods and hopes in the future to be able to narrow the number of different types of analyses Even with the multiple analyses Pipberger's computer program is able to analyze 700 to 800 records per hour

Young and Huggins are using a program which determines orthonormal exponentials allowing each electrocardiographic lead to be represented by twelve numbers Since there are three orthogonal leads this makes a total of thirty-six numbers Nine additional terms are used to represent the direction cosines that describe the rotation of the heart in space, making a forty-five dimensional vector Standard multiple dimension linear regression analysis is then applied and each subject is assigned an additional attribute vector according to the diagnostic category For example, the attribute vector of a normal subject is (1,0,0,0,0,0) Other groups of patients having left ventricular hypertrophy, right ventricular hypertrophy, left ventricular conduction defect, right ventricular conduction defect and infarction were each assigned unique attribute vectors The categorization of disease was carried out by a physician It was possible to accomplish perfect diagnoses in the fifty-three subjects studied by this program, but the number of patients studied was admittedly inadequate to judge the usefulness of this approach

Enslein has also recently suggested the application of pattern recognition methods to biomedical problems, particularly the interpretation of x-rays and electrophoretic patterns, but also the reading of electrocardiograms This approach is not too dissimilar from that of Young and Huggins The only requirement as far as the analysis is concerned is that the program be trained by *assigning arbitrary numbers to each of the disease classifications* and then that an appropriate number of known cases of each of the diseases be *shown to the computer so that when that pattern* is later shown as an unknown the various exponentials will be correct to give the designated answer Thus, after training of the program, right ventricular hypertrophy might have been given an arbitrary number five, normal electrocardiograms the number one, etc and then when an unknown is shown if it is normal the number one should show up as the answer It has just been announced by one of the computer companies that they have developed a scanner which will identify printed letters and numbers and which is called a Multi font optical character reader It was observed during study of the scanner that the computer linked with this device performs some functions that have not been worked out completely The early scanning rate is fifty characters a second, but the potential speed is hundreds of characters a second It is hoped that this approach and type of equipment can be used *in language translation and there is no reason* why the translation of Russian to English should be any more difficult for pattern recognition techniques than the translation of electrocardiographic inscriptions to printed diagnoses

It is very obvious in the pattern recognition approach to computer analysis of electrocardiographic tracings that the physician's chief duty is to provide very accurate diagnoses of the patients whose tracings are used for training the program This is equally true of the other analytical methods, but it becomes less obvious when the computer is programed to simply carry out the standard analytical techniques that are now in use Validation with the necessary statistical analysis of results comes from careful follow up of patients and on going revisions of diagnoses and corrections of the computer programs in use as the experience accumulates There have been studies of large numbers of supposedly healthy young men which showed relatively large numbers of interesting and supposedly important electrocardiographic findings, but much of the value of some of these studies is lost in that no long time follow up has been carried out

If all a physician wishes from a computer is an electrocardiographic based diagnosis then there will be programs available to provide this with the accompanying probability plus a list of differential diagnoses with their respective probabilities This will not prevent the physician from studying the electrocardiographic tracing Schmitt has emphasized the additional use of computers in combination with human visual perception and possibly even accompanied by a loud speaker output so that the physician may listen to certain parameters of the tracing while he watches a stereoscopic view of the vector loop in three dimensional space Schmitt has developed ways in which a stereoscopic projection can be rotated and viewed from different directions in three dimensional space with a cathode ray oscilloscope

Abildskov, Ingerson and Hisey call attention to the deficiency of vector cardiograms as recorded on oscilloscopes as far as the temporal relationships are concerned and have developed a spherical co ordinate system which describes a point on the surface of a sphere The vector is then presented in terms of magnitude and two angles Such displays might be helpful for study of selected patients with suspected posterior myocardial infarction Perhaps some of these visual techniques can be used to work out the problems that arise when heuristically programed computer programs for pattern recognition come up with information that cannot be classified It seems as though more attention should be given, at least in some laboratories, to the on line analysis of electrocardiograms in the fashion of the electroencephalographers (even though the problems of electrocardiology may be "trivial" by comparison) No doubt the electrocardiographic laboratory of the near future will contain equipment to record phonocardiograms, peripheral pulse contours, respiratory patterns, respiratory gas analyses and other information all simultaneously and there seems to be no limit to the computing that could be done

The electrocardiographic tracing like brain waves serves as a source of biologically meaningful signal which is used in system feedback as a transponder function This is detailed in the next chapter, but brief mention will be made here Once there is identification of a segment of the tracing such as the R-wave then this can be used as a timing signal to activate peripheral equipment which is required to function in a particular phase of the cardiac cycle In patients with recent coronary artery thrombosis and myocardial infarction or in patients with hemorrhagic, traumatic or bacteremic shock, who have developed a significant component of pump failure and need assistance with their circulation until the heart can recover, it may be possible to assist with a variety of mechanical devices Thus a patient may receive "artificial circulation" just as he may receive assistance by "artificial respiration"

Claus and recently Harken and Lefemine have investigated the use of a large cannula and aspiration of blood from a large artery during systole (while the heart is pumping blood out into the arteries) and then blood is pumped back into the same artery during cardiac diastole (that period when the heart is relaxed or not pumping) Thus the heart pumps against a very low resistance system and additional propulsion of blood is provided from an outside source of energy Others are using an intraaortic counter-pulsating balloon The most appealing idea for ease of application and safety if it can be developed is to use either the lower extremities or the intrathoracic contents as a second heart

Osburn, Russi, Sale, Bramson and Gerbode have shown that both methods are possible Both are timed with the electrocardiogram A pressure suit is used for the legs and the suit is pressured during diastole of the patient's heart so that the patient's diastolic blood pressure is actually raised above systolic and cardiac output in dogs has been increased by 30 per cent When the pulsation is used in the airway the timing is regulated so that the external pulse rise is synchronous with cardiac systole and the endotracheal pressure is dropped during diastole with no change in mean airway pressure and no interference with the normal respiratory rhythm

Chapter V

MEDICAL RESEARCH

WHEN should computer use be considered? Automation in research should not be undertaken without good reason The initial steps are prodigiously expensive if equipment is not already available and always consume sufficient time and energy as to make the results seem illusory for the small one-run problem and often diversionary in relation to the original question These objections are minimal if the problem involves use of standard, already available equipment and programs Even the experienced user may hesitate to automate a new project because of the time and expense required in preparation of data The reasons for automation might be divided into three general categories (1) The volume of data requiring storage, classification, arrangement into various relationships, and the number of simple analytical procedures, have become humanly impossible (2) The complexity of the analyses, the number of variables, the number of inter-related equations makes the study impractical without machine processing, and (3) The speed with which the results of analysis are required exceeds human capabilities

There are also degrees of automation in research varying from the use of a computer much as one might use a desk calculator to carrying out the entire experiment within the computer Examination of the varying degrees of automation in research suggests a classification which will be used in presenting a number of otherwise unrelated research projects and these subdivisions will be taken up in the following order (1) data housekeeping and routine statistical analyses, (2) on-line computer investigator feedback, (3) transponder function, (4) dynamic simulation, (5) adynamic simulation and (6) combinations of the above An attempt will be made to define these different uses, but subsequent examples should help also and the reader should not give up on the next few pages First, we need some definitions of terms *On-line computation* means that the computer takes the data in real time and carries out some or all of the necessary computations immediately so that the investigator may have use of the results in the experiment while that experiment is being run The investigator can, therefore, use the computer results to modify experimental procedure A *transponder function* is programed ahead of time so that during the experiment the computer not only analyzes some of the data and prepares the results of the analysis for use in the experiment, but automatically responds to certain predetermined circumstances that arise in the experiment and provides appropriate modification of the experimental procedure without any interference on the part of the investigator

In *dynamic simulation* the computer analysis is carried out in real time with components designed to simulate the system under study A signal is taken from the living subject or provided for both the living subject and the computer The signal is processed in its normal manner by the living organism and also by the simulation program in the computer and then the observed and calculated results are graphed so that the investigator may check on the accuracy of his model Dynamic simulation is a real time experimental situation in which the investigator can immediately alter various coefficients or components in the model with the living experimental subject acting as a very close check on his model building

With a transponder function the computer becomes a component in the experiment In dynamic simulation the computer and the experimental subject may be completely separate as far as any automatic feedback is concerned Dynamic simulation is more than just computer analysis of experimental data since the objective is to have the computer simulate the actions and reactions of the experimental subject The investigator makes adjustments of the computer to make it a better simulation Since the experiment is in real time and since the experimenter wishes to have a model which follows the living subject over as wide a range as possible, the ongoing results will govern some of the sequence and parameters chosen for variation by the investigator

Dynamic simulation is modeling with analog and experimental subject in parallel and with only qualitative or rough quantitave preconceived information as to the functional characteristics of components involved Continued real time human inspection is used as a substitute for more elaborate theory and the subsidiary hypotheses as to the actual quantitative and mathematical relationships that will be found Real time simply means that the computation is completed immediately and results are made available so that the real and calculated signals can be compared in the same recording

Adynamic simulation is modeling in sequence, off-line, at any time, *in vitro* and completely within the programed computer It requires a complete physiologic, hydraulic, or metabolic, map with every stationary and kinetic component predetermined within acceptable limits Predictions are made by use of the program and are then checked by specific experiments and chemical or other analyses for intermediary products or signals in accessible points of the living system A variety of medical research projects will be outlined to illustrate in greater detail the variation of computer use and at the same time some of the methodologic problems

DATA HOUSEKEEPING

The first example illustrates a research project which was started by the National Institute of Health on January 1, 1950 and has led to the use of computers by virtue of the large volume of data which must be processed The type of equipment and methods of use of that equipment are essentially those used in business Originally the data were put on standard punch cards and then as this became too cumbersome and with more recent availability of magnetic tape storage, the information has been transferred to tape

This was begun as a prospective epidemiologic study of heart disease in Framingham, Massachusetts, a town of 30,000 population, eighteen miles west of Boston An initial random sample involving two-thirds of the residents between ages thirty and fifty nine was drawn and these individuals were then invited to participate in the study Only 68 8 per cent of these people actually came for examination Some 740 volunteers were then added to the original selection so that the total group who were found to be free of coronary heart disease at the outset was 5,127 An attempt was made to find out as much as possible about the individuals in the community who were not included in the study, as far as their mortality rate and development of known cardiac disease was concerned

The individuals included in the study have been subjected to thorough initial examinations and repeat examinations at two year intervals, consisting of history, physical examination and routine laboratory tests with special studies such as electrocardiograms, blood cholesterol levels and chest x-rays for all participants and more detailed analyses as indicated This would seem to be *a nearly ideal experiment, but no experiment can ever be completely prospective and perfect in design*

The early death rates in the participants and in the town as a whole were difficult to assess because of lack of at least minimal data on the health status of the entire population prior to the random selection of participants It is not known exactly what the relationship was between state of health and refusal rate of the residents Another question which has been hard for the investigators to evaluate (and is often present to a degree) is the effect of the experiment per se (the fact that the experiment is being done) upon the results Whenever any abnormality is found, the patient's family physician is notified and further investigation and possible treatment is instituted which could be another source of bias

These investigators have also made note of the concern that some statisticians have for using statistical procedures in investigations where subgroups are selected from the original population on the basis of certain criteria instead of selection in a random manner Gordon, Moore, Shurtleff and Dawber observe "A statistical statement must begin and must end with the probability model Thus when you compare the mean blood pressures for a coronary and a noncoronary group the probability statement you make refers to two sets of numbers, assumed to conform to a specified probability behavior, not to blood pressure or coronary disease You may decide the means of the two sets of numbers are or are not the same by whatever statistical rules you have chosen to use This is a statistical decision When you go beyond this to make assertions about coronary heart disease and blood pressure the inference may or may not be justified, but it is not statistical "

One of the interesting questions about experimental design which this study helps to answer and which may help in the validation of other studies of selected groups of individuals with coronary disease has to do with the question of the selective effect of the disease event or of death Looking at this from the other direction, if an investigator sets out to study coronary artery disease and gathers together all of the hospital charts of patients with this diagnosis, does he have a representative sample of all patients with coronary disease or does his sample have considerable bias because of the patients already eliminated by death or eliminated by the fact that their coronary artery disease has not yet been recognized? It was observed in the Framingham study that serum cholesterol levels, blood pressure and weight were essentially the same after recovery as they had been in these same patients before Furthermore, the weights, blood pressure and cholesterol levels for men who developed new coronary heart disease were the same for the men who died as for those who survived So far nothing has been found in this study which should discourage retrospective study of coronary heart disease

In a massive study of data of this sort availability of computer equipment encourages a much more thorough study of many more relationships, but also helps in studying the methodology per se, and at times points up some of the inadvertent bias that human participation contributes For example, a study was made of the distribution of reported blood pressures according to the last digit in the blood pressure It was observed that there was a predilection for recording even digits and the most popular observation was the end digit "zero " Sixteen hundred and fifty-one systolic blood pressures were observed to end in "zero", as contrast to only one ending in the digit "one," and only one ending in the digit "three " For diastolic blood pressure there were 823 observations ending in "zero" and none ending in the digit "one " The digits "two," "four," "six" and "eight" seemed to be very popular and the odd digits unpopular except of "five" It would be interesting to observe the difference between blood pressures recorded by human observers and by a mechanical monitor in such a study

Computers can assist in a large study of this sort not only in examining the distribution of data, but in standardizing the scales used At first there were, for example, only the initial values of cholesterol for each patient, but with each successive examination, additional cholesterol values have been added to the over-all study In order to use cholesterol in comparison with other measurements it was desirable to standardize the values by converting them all to standard deviations This, of course, required computing a mean for all cholesterols measured prior to the time of the computation, determining the deviation of each result from this mean and squaring all these differences They are then summed and the root mean square deviation, or standard deviation, was then available so that the computer could convert every measurement from its original scale of milligrams per cent to standard deviations Furthermore, the scale can be accurately divided into percentiles and each cholesterol level can be categorized as to its percentile rank This makes possible a more definable study of the effects of cholesterol on the incidence of disease This type of standardization of scores is of even greater importance when the interactions of several variables are under examination This will be gone into in some detail in Chapter VI

Large intensive studies such as the Framingham study are to interest not only to students of cardiovascular disease, but of problem solvers in general who must use similar methodology There is continued self criticism of various aspects of the study as well as stimulating criticism such as the paper of Lew in which he discusses the effects of persons refusing to co-operate in the study and other forms of self selection and the difficulties of interpretation and generalizations from the observations that are made This large study shares with smaller studies the need for repetition in other geographic areas and under other circumstances to confirm the observations made and to determine to what extent generalization is justified from the results

A computer oriented investigation is of necessity organized in such a way that data can be selected and compared or combined with data from other studies New retrospective investigations can be made with minimal additional work involved since the data has not only been collected but is accessible An example of this is the recent publication of Doyle, Dawber, Kannel, Heslin and Kahn on cigarette smoking and coronary heart disease in which they combined information from the Albany Cardiovascular Health Center and the Framingham Study This gave them a total of 4,120 men who were free of coronary heart disease at their first examination The results of this analysis revealed that heavy cigarette smokers experienced a three fold increase in the incidence of myocardial infarction and in death from all causes as compared to nonsmokers, pipe and cigar smokers and former cigarette smokers These findings, according to Lew are based on more reliable statistics than many of the other findings in the Framingham Study, but how the results are interpreted has nothing to do with computer use

The computer is a tool which makes possible the examination of data from many viewpoints and makes possible studies which otherwise would not be undertaken or would be undertaken on a much more restricted scale Certainly computers will make the concept of preventive medicine much more important by facilitating long-term supervision of initially asymptomatic adults in order to delineate those characteristics which are most likely to be associated with high risk of certain diseases The practicing physician will be better able to identify susceptible individuals He will have a better means of estimating the magnitude of the risk It may even, we might hope, be possible to diagnose and counteract epidemics of disease, even chronic disease, if the pertinent information can be obtained and appropriate counteracting measures discovered

It appears so far from the Framingham Study that cholesterol is a better predictor of coronary heart disease than other lipids that were measured Systolic blood pressure is a valuable predictor Whether or not there is true interaction between some of these factors, as far as prediction is concerned, a combination of two or three of the abnormal variables studied, becomes quite highly predictive of overt coronary disease

It is not necessary to institute a large scale study and to collect thousands of multivariant observations in order to find applications of computers If there is a computer center available, standard library programs for a variety of statistical analyses are ready for use For example, Hidalgo, Nadler and Bloch, used an electronic digital computer and a regression analysis program to determine, from data collected on some ninety-two men and sixty-three women, the best fit for the relationship of blood volume to height and body weight The regression analysis program will be discussed further in Chapter VI

The most intriguing applications of computers are in problems which are too complex for man's limited capacities of storage and association Dayhoff and Ledley have developed a series of seven programs for use in protein biochemistry Proteins are long chain polymers of amino acid molecules There are only about twenty amino acids, but each protein is different because of the order in which these acids are arranged All enzymes and cell structures are dependent upon the arrangement or the coding of amino acids in protein The amino acids in turn are produced on a template of nucleic acid material which is also coded A real understanding of life, inheritance, evolution, growth, development and cellular metabolism sooner or later leads back to the question of how these protein molecules are put together Because of the coding problem, the study of protein structure is even of interest to the computer industry

Ordinarily chemists have laboriously and rigorously analyzed chemical compounds by determining the atomic arrangement of each part and by successive degradations and chemical reactions With the development of x-ray defraction analysis and the appearance of electronic calculating machines it was possible to make more rapid progress Pauling has described the stochastic method by which organic chemists have studied structure and in which the chemist makes an educated guess as to the structure on the basis of the information at hand and then proceeds to carry out a few experiments to check specific points regarding his conjecture This has been difficult with the protein molecules because of their size The original protein molecule is usually broken down by enzymatic hydrolysis into shorter chain structures of amino acids called peptide fragments and then the chemist must conjecture as to the sequence of these fragments Some of the protein molecules in a sample are broken at different points in the chain from other molecules so that there is an overlap in the peptide fragments There are certain amino acids which occur infrequently or in unusual relationships

Dayhoff and Ledley constructed a model of a protein using different colored "pop" beads to represent amino acids The exact arrangements of the "amino acids" in the model were recorded and then the model was randomly fragmented into peptide-like fragments and these fragments were also listed The data derived was then used to test the series of programs which were written to determine the possible ways in which the original molecule could have been arranged to give rise to the fragments observed The program took into account the possibility of some errors in peptide analysis and ways in which these errors could be eliminated The program was written so that the input and output would be alphabetic and mnemonic Three letter abbreviations were used for each amino acid Intermediate results of the analysis are printed out so that the biochemist can study the sequence of synthesis that the probability theory program is going through in order to reassemble the molecule It is possible to insert special rules into the program regarding certain end of the chain groups that are known to exist for the protein molecule by virtue of studies of the first amino acids that come off during hydrolysis

The programs are thus written so that the biochemist can and does actively participate in the computer analysis Just as the biochemist has a very elaborate amino acid analyzer to help determine the fragments as they come apart during hydrolysis he may use the computer as another tool for the theoretical resynthesis, but the chemist remains a very important part of the system and it is unlikely that the structure of proteins could be determined through computer use by a completely automated program

ON-LINE COMPUTING

On-line real time computation with the investigator making immediate use of the results to further modify the experiment has been discussed in the previous chapter in relationship to electroencephalographic special purpose computers Zimmer has described this application in a psychophysiologic study of autonomic nervous system reactions The autonomic nervous system is, as its name implies, part of the more or less automatic regulatory system of the body which provides homeostasis and it is affected by emotional states Zimmer records respiration, heart rate, peripheral circulation, skin conductance, muscle action potential, body movement and reaction time This information is recorded along with data concerning the time and duration of stimulus presentation and the onset and termination of re enforcement symbols

A computer program takes the nearly 9000 bits of data which are collected in the course of thirty seconds following a single stimulus and summarizes this in what is called a Lacey's Autonomic Lability Score This is recorded as an index of the subject's reaction and is related to the prestimulus baseline This sounds a little bit like what the lay person thinks of as a lie detector Experiments of this character might reveal the least number of types of observation which might be useful in monitoring a patient's emotional stability and response while he took a Cornell Medical Index and a Minnesota Multiphasic Personality Inventory, but away from "cloud nine" and back to the illustrations

TRANSPONDER FUNCTIONS

It was mentioned in the last chapter that respiration had been observed by Einthoven and fellow workers to influence the electrocardiographic tracing The elevation of the diaphragm lifts the heart and rotates it with each respiration which has an especially marked effect on spatial vector cardiograms If there are three types of signals mixed together from the conduction system of the heart, from the effect of change in position of the diaphragm, and from interfering noise of sixty cycle current and muscle tremor, and if the relationships between breathing and cardiac cycle are irregular, it becomes very difficult to obtain data which allows analysis of the effects of respiration on the electrocardiographic tracing The transponder function solves this problem by keeping track of the electrocardiographic wave sequence and signaling the patient when to breathe and also timing the recording of the electrocardiogram so that repeated tracings are superimposed and averaged As each succeeding curve is averaged into the previous one the pattern becomes progressively clearer as the unwanted noise is averaged out

Transponder functions are important considerations in the development of patient monitoring since they may be used to eliminate the recording of large volumes of normal data, but can be set to institute recording as soon as abnormalities appear Schmitt has suggested that transponders be considered for appropriate institution of treatment under certain circumstances One might consider, for example, the addition of a transponder function to the cardiac pacemaker in a patient with intermittent complete heart block so that as long as the patient's cardiac rhythm was at an approximately normal rate the pacemaker would be inactive but as soon as the normal rhythm was interrupted the cardiac pacemaker would take over

DYNAMIC SIMULATION

Warner has used an electric analog of the carotid sinus in which the computer acted also as a component in the experiment, taking recordings of arterial pressure from the carotid sinus and feeding additional stimulation back into the vagus nerve so as to simulate and exaggerate carotid sinus regulatory function They found that the carotid sinus was important in determing steady state levels of mean blood pressure, but did not respond to rapid fluctuations in pressure

Clynes has studied the respiratory heart rate reflex and its change with age and has suggested that permanent records could be maintained for individuals, collecting additional data over the years and watching for the normal changes that occur with aging as well as unexpected changes which might herald the development of some disease process at a stage before it would otherwise be recognized

The eye has proven to be a very common subject for study of biological servo systems Stark has studied the pupil area by measuring the reflection of invisible infra red light with changes in intensity of visible light as the stimulus The ciliary musculature is a very accessible and responsive muscle to observe with only one degree of freedom in its motion Considerable is known about the anatomy of the pupil and reflex which involves the light path, optic nerve, the lateral geniculate body and some fibers which go directly to the brain stem and are relayed from there to the Edinger-Westphal nucleus which in turn contains parasympathetic nerve cells whose fibers can be traced to the ciliary ganglion which then controls the iris muscle There are other nerve cells which send fibers through the sympathetic nerve system through the superior cervical ganglion and cause dilatation of the pupils

Three varieties of light stimulation are used When a wide beam *is shone through the pupil the amount of light reaching the retina* is decreased as the pupil constricts which is called a closed loop type of stimulation since, as the pupil constricts, there is less light hitting the retina and, therefore, less stimulation to cause contraction of the muscle This is the normal closed loop feedback mechanism A second type of stimulus (open loop) is to shine a very narrow beam of light so that, although the pupil is stimulated to constrict, this constriction causes no decrease in the amount of light reaching the retina A third type of stimulus is to shine a narrow beam so that it impinges on the edge of the pupil and very slight contraction of the pupil will have a very marked effect on the amount of light passing into the eye Thus there is a very high gain effect in the feedback mechanism By use of an appropriate stimulus the regulatory system for control of light entering the eye may be studied as a normal closed feedback, an open loop with no feedback, or with high gain (exaggerated) feedback Other workers have used computer techniques to study the servomechanism controlling eye movement in tracking a moving object or in maintaining a fixation upon an object while the head and body are moving

SIMULATION

Simulation need not be on-line It can be completely *in vitro* with separate laboratory living system experiments done later to test the results predicted by the computer program Simulation offers some of the greatest challenges in computer applicatio since it offers a means of testing hypotheses about the inte relationships of system components as well as study of the components themselves Warner and Cox used analog computer simulation of the mathematical model of the heart rate control b sympathetic and vagus nerve function and compared the computer model with anesthetized dogs In building their model the took into consideration information gathered from other investigations regarding the concentration of norepinephrine in it storage sites in nonmedulated sympathetic nerve fibers, certai characteristics of its release, the concentration of circulatin norepinephrine in the blood, certain characteristics of the sino auricular node in the heart which controls heart rate and postulated substance in the node which theoretically must reac with norepinephrine in order to change heart rate This involve postulation of some six different rate constants The model predicted that raising the concentration of norepinephrine in circulating blood would have no effect on the maximum heart rat achieved by sympathetic stimulation, which was confirmed in th experimental animal They also set up a similar model of th vagus nerve function and its discharge of acetyl choline Thei model seemed to account for simple stimulation of either th sympathetic or vagus nerve, but could not account for effects observed when both forms of stimulation were used

Brittain Chance has written digital computer programs containing equations for all known steps in the chemical reactions that take place for selected segments of cellular metabolism This allows experimentation with some of the less accessible portions of the simulated living cell and then when actual experiments are run with the biological system, only spot checks need to be made to test the accuracy of the computer prediction

There are close similarities between the models used for enzyme kinetics, studies of transport of compounds between various compartments of the body and for epidemiology An enzyme is often thought of as a specific type of small ferry boat which plies back and forth across the energy barrier of a specific chemical reaction, picking up substrate on one side and depositing the passenger molecules as product of the reaction on the other side on the opposite shore Competitive inhibitor compounds of similar structure to the substrate may also sit in the seats and thus slow the rate of transport of the substrate The same mathematical models are applied to the transport of chemical compounds across the kidney tubule and across cellular membranes throughout nature

In epidemiology the ferry boat is not a single enzyme or a specific transport substance but is the bacterial agent and the passengers are people who have not previously had the disease They become infected and are carried across the river to the immunity community (or across the river Styx), but the mathematical models and computer programing have much in common A well trained and interested mathematician could find ample problems in biology and medicine for exercising his talents All biologists must make an effort to cross this barrier too, but this is a subject for the last chapter

Chapter VI

A WORKED EXAMPLE

THE clinical investigator or medical research worker who wishes to use automatic data processing facilities usually feels initially somewhat like a child waiting for his parents to begin a vacation trip There seem to be so many preliminaries which delay actually getting started The final computer data processing may take only five or ten minutes The development of hypotheses, design of experiment, collection of data and preparation of data for analysis may require months or years Clinicians often have their enthusiasm dampened further by the rigor of the approach required by automatic data processing techniques

The work need not necessarily be mathematical, but it is at the least very logical and requires most careful definition of each step The research worker who has mastered clinical medicine and required portions of chemistry, physics, physiology, anatomy, etc may suddenly realize that he needs much more knowledge of statistics There are specialists in related disciplines who can provide a certain amount of help, but the only specialist in that particular research project is the investigator himself and just as he needs to know a great deal about all of the other aspects of his experiment, he must also have considerable insight into the logical, statistical, and mathematical aspects of his work

A narrative review of a research project which required the use of automatic data processing has been chosen for more detailed illustration of the application of computers since it involves the types of problems that are frequently met with in clinical research There will be frequent interruptions in this account in order to explain various statistical matters or to illustrate from earlier experiments why certain analyses were considered desirable or to explain aspects of actual computer programing and the details of computer use in this study First, a short resume of the research problem will be presented

Our currently complex treatment of patients with a variety of illnesses and with extensive surgical procedures has saved many lives but has brought out the problem of acute renal failure which often adds to the patient's illness and further complicates the treatment There are at present no available means of adequately evaluating the diagnosis and prognosis of acute kidney damage at a time when something specific could be done to reverse the process There are certain drugs that are occasionally used in the treatment of resistant infections which are known to be toxic to the kidney In spite of careful crossmatching of blood there are still occasional transfusion reactions followed by kidney failure There are also a variety of acute and chronic diseases of the kidney which are recognized primarily in their late stages by discovering the rising blood urea nitrogen

In all of these instances a more sensitive means of diagnosing and following active early damage in the kidney would be of considerable benefit to the clinician Not only would early diagnosis be a help but it would be possible then to evaluate the success or failure of a proposed treatment For example, hundreds of patients are being given intravenous Mannitol at the present time during and immediately after surgery to prevent renal failure, but there has been no specific way to estimate which patients might have had renal failure had the Mannitol not been given

Serum enzyme concentrations are used as indices of organ damage, as for example in the study of glutamic oxaloacetic transaminase as an index of acute liver cell breakdown or acute myocardial infarction The basis for these analyses rests in our concept of enzymes as being large protein molecules residing in cells and being released from cells into the blood stream or surrounding tissue fluids only when those cells are damaged In order for such tests to be useful the blood or tissue fluid sampled must have come into contact with the cells under question In acute damage to the kidney there should be loss of cellular enzymes into the urine

The overall, kidney, enzyme losing, experimental model could be, of course, much more complicated than this Almost 200 quarts of fluid are filtered from the blood each day by the kidney glomeruli and all but one quart reabsorbed by the tubules The blood cells and larger protein molecules in the blood are prevented from being filtered through the glomerulus into the nephron by virtue of the small size of the pores in the glomerular filter Some smaller protein molecules may slip through, especially if there is damage to the glomerulus

When protein does pass into the nephron it is phagocytized by cells lining the tubule and much of it is probably removed from the fluid which eventually becomes urine Certain substances are excreted by the cells lining the nephron and enter the urine by this active contribution of cells rather than by passive filtration through the glomerulus It is probable that in the course of the excretion of certain solids into the lumen of the nephron some protein molecules may be lost in small amounts

The amount of enzymatic protein which reaches the urine may be affected by the rate at which the fluid passes through the nephron, thus altering the time available for removal of a given enzyme molecule by phagocytosis Also if the tubule cells are required to do an excessive amount of osmotic or chemical work there may be additional loss of enzymes The enzyme content of the cells is qualitatively different at different levels of the nephron and the pattern of enzyme excretion by the kidney could very well be related to the functional activity of the proximal and distal convoluted tubules and the ascending loop of Henle as well as the collecting tubules Enzymes could also be lost from the collecting system of kidney pelvis, ureter, bladder and urethra There may be a certain obligatory loss of enzymes as cells deteriorate and die and are replaced by new cells This might be difficult to distinguish from mild toxic effects

Initially the urines for this study were obtained from a variety of normal ambulatory individuals as well as normal hospitalized patients and patients with potential kidney disease It was found that certain enzymes were stable in urine and could be measured with an accuracy sufficient to determine that there was a difference between hospitalized patients and normal, well, ambulatory people Certain patients were studied in detail, collecting serial samples of urine and it was observed that there were curvilinear variations with time which were not random and must, therefore, have some significance either relative to kidney function or to some aspect of the integrity of the kidney

The urine enzyme levels did not seem to be related to the serum enzyme concentrations, suggesting that the enzymes in the urine came from the kidney rather than from the blood and other organs in the body Three enzymes have been studied extensively, namely glutamic oxaloacetic transaminase (GOT), beta-glucuronidase (B-G) and leucine amino-peptidase (LAP) A series of experiments were carefully designed using dogs in metabolism cages and varying the intake of food, water and solute in some of the experiments and, in other experiments, adding agents which are considered to be toxic to the kidney

In one experiment where both Neomycin and egg albumin were injected intravenously, there appeared to be a highly significant increase in the loss of all three enzymes in the urine With the injection of egg albumin alone the rise in enzyme excretion was less striking There was additional question about the results of the experiment because of the occasional need to discard urine specimens contaminated in the metabolism cages from nonformed stools immediately following administration of the toxic agents A new experiment was, therefore, carried out and will be used for illustration of computer applications below

It is known that egg albumin introduced into the circulation as a foreign protein causes the development of an allergic reaction in animals and also that the foreign protein can be found in the tubules of the kidney The experimental design chosen called for *injection of egg albumin on Monday mornings and urine collections* beginning at 0800 hours after the injection Twenty-four hour urine collections were then made on Tuesday, Wednesday, Thursday and Friday mornings at 0800 hours If any emesis or diarrhea occurred immediately following the injection, the cage was cleaned before urine collection was started

In addition the dogs were all maintained on a reduced dietary intake Friday, Saturday and Sunday and until noon on Monday The diet consisted of kennel biscuits containing 448 calories and 4 05 grams nitrogen/100 grams biscuit The feedings were divided into two equal feedings, one given at noon and the other at 4 o'clock in the afternoon The food was weighed out and given according to dog weight Earlier it was determined that the average dog in a metabolism cage would eat a certain amount of this diet when the diet was provided in unlimited quantities In the present experiment one-third of this ad libitum average amount was given to all of the dogs on week ends This greatly reduced the incidence of vomiting and diarrhea following injection of egg albumin It did, however, introduce another variable

Earlier experiments had shown that variation in food intake affected markedly the excretion of leucine amino peptidase (LAP) and to a lesser extent the excretion of other enzymes The present experiment was therefore designed to evaluate the effect of variation in food intake In order to do this the dogs were divided into two "C" groups Both groups were fed the one-third ration over each week end and group "two" was fed the one-third ration during the week as well Group "one" dogs were given three-fourths of an average ad libitum intake on Tuesday through Friday The general layout of the experiment is shown in Figure 1

The main simple effects to be observed and outlined thus far in the experiment were therefore defined as (A) the influence of injected egg albumin and (C) the amount of diet There could also be some interaction between A and C For example, the reduction in diet in group "two" animals (C2) could make the kidney more or less susceptible to the effects of a foreign protein as measured by excretion of enzymes in the urine

Another condition seemed to be in need of consideration in the experimental design, namely the influence of repeated injections of foreign protein as it might affect the sensitivity of the animal This might be manifest as an increasing or decreasing loss of enzymes after injection of egg albumin as the kidneys became more or less sensitive to the foreign protein There was also a possibility of another interaction or combination effect For example, the group "one" dogs (C1) receiving the higher food intake during the week might have more enzyme to lose from the kidney, but this effect might become less with time Thus there could be a three way interaction Interactions might also be observed between varying sensitivity and diet alone, or between the effect of egg albumin injection and the sequence of injections In order

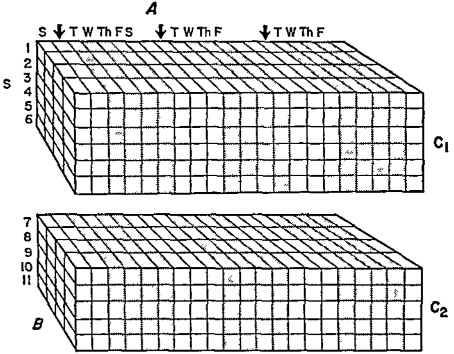


Figure 1 This is a Lindquist Type VI analysis of variance design It can be considered a four dimension table in which a criterion variable such as the urine concentration of leucine amino peptidase (LAP) could be entered for analysis It is apparent that variation can be analyzed by dogs (S), days (A), treatment repetitions (B) or differences in diet for groups of dogs (C) Arrows show injection times for all dogs in both C_1 and C_2 Shaded areas show the actual days when urine samples were collected Table II shows the results of a Type VI analysis for LAP Figure 4 shows the mean values obtained from the analysis Similar analyses of variance were carried out using other urine constituents as the criterion variable but the results are shown only for LAP The analysis is based on the null hypothesis that there is no significant difference in the variation of the criterion variable which can be attributed to A, B C, AB, AC BC or ABC influence

to provide a factor of "influence of *repeated* exposure to the foreign protein" the experiment was replicated four times along another dimension which will be designated as "B" The simple effects A and B and the interaction effects of AB, AC, BC and ABC are all effects in which comparison can be made *within* dogs and the difference between animals does not need to be considered as affecting the analysis Each dog provides both control and experimental data The C effect or difference between levels of feeding *is of necessity a **between** subjects effect Dogs are either in C1 or* C2 and not in both

A standard *mixed* design was selected which provided this combination of *within* dogs and *between* dogs effects by consulting a summary Table [# 17, Lindquist] in a standard design and analysis text This particular analysis is called a Type VI design In the State University of Iowa Computer Center, analysis of variance programs, written by Spiker and compatible with Lindquist's text are available on a magnetic tape in the computer program library It is usually possible to select a standard design which will fit the experiment This has not only the advantage of saving a great deal of time for the investigator, but such standard library programs have been tested not only by their originators but by many different users With the experimental problem and a specific design outlined, the next step is to do the experiment

Healthy mongrel dogs were randomly selected and placed in metabolism cages and provided with the diet as it was to be given *during the actual experiment In addition egg albumin was given* intravenously in a dose of 1 ml/kg on each Monday just as if the experiment had been started, but no urine samples were collected The dogs were thus sensitized to egg albumin before the actual experiment was begun After three weeks of allowing the dogs to become accustomed to the routine, urine was then collected from the eleven dogs in the two C groups for four days of each week and for three weeks making 132 samples This was called a B slice and the entire experiment was then repeated three more times so that 528 urine samples were obtained in four B slices for the entire experiment An estimate of the average value for any *measurement made in these samples would, therefore,* have 527 chances to vary, or 527 degrees of freedom

It would seem somewhat wasteful to spend three or four months collecting urine specimens and making the injections for the purpose of answering only one question Actually three enzymes were studied in the urine so that an analysis of variance could be made for each of these three enzymes Earlier experiments indicated that LAP would be the most likely enzyme to be influenced by diet Earlier experiments *with the injection of nephrotoxic agents also* revealed that all three enzymes were elevated in kidney damage

Other urine analyses were carried out in order that multiple regression analysis could be used For example, in one earlier experiment which lasted 130 days, a group of thirteen dogs were given a varying diet, variation in water intake and variation in the sodium content of the water Multiple regression analysis then revealed that when GOT was the dependent variable and creatinine, sodium, potassium, chloride, phosphorus and freezing point depression were used as independent variables, that potassium, freezing point and sodium were all important predictors of GOT excretion, both with concentrations of these moieties and with total amounts excreted per day The partial-F for potassium was especially high The correlation between potassium and GOT excretion was positive and between GOT and sodium negative This suggested a hypothesis that perhaps some of the GOT in the urine is lost from the distal convoluted tubule coincident with the exchange of sodium for potassium

In other earlier experiments where sodium intake was low throughout, there was no correlation between potassium and GOT excretion This would be consistent with the proposed hypothesis since it is known that potassium in urine remains at a lower level as long as sodium reaches the distal convoluted tubule in relatively low concentrations This does not prove the suggested hypothesis, but does illustrate the way in which examination of the data may reveal other interesting and worthwhile avenues for investigation Multiple regression analysis also has the potential of revealing certain relationships between those criterion variables selected for analysis of variance and other variables measured in the same urine samples Results of regression analysis may assist in explaining the source of the effect of those significant factors in the experiment such as diet or the injection of a toxic agent If significant relationships are found between the excretion of an enzyme and the excretion of other moieties, then covariance analysis can be used to further delineate the relationships and may point the way to additional experiments which may explain in greater detail the reasons for the effects observed in the analysis of variance

The present experiment was, therefore, elaborated to include measurement of these other moieties, urine volume, pH, specific gravity, total solute (by freezing point depression), creatinine, sodium, potassium, chloride and phosphorus concentration, in addition to the activity of three enzymes GOT, LAP and B G After the entire experiment had been run, specimens collected and analyzed, data recorded in the laboratory book and also on punched cards and ultimately upon magnetic tape it would seem that surely now some computer analyses could be started

Before reporting upon the actual outcome of the analysis, it is worth mentioning some of the underlying assumptions that must be fulfilled or at least considered before carrying out the actual computation

The analysis of variance is a very useful statistical model, but in considering this model certain conditions must be met if the results are to be interpreted The first assumption is that the criterion measure for each treatment population is "normal" or Gaussian in distribution The necessity for adherence to a precise model cannot be over-emphasized Many of the models of mathematical statistics dealing with the mean and standard deviation are valid only when the population has a normal distribution Secondly, the analysis of variance is based upon certain standard hypotheses (1) all groups were drawn from the same population, (2) there is no difference in the population variation in the different groups, and (3) the population means for each group are the same If any of the above conditions are not true then the analysis of variance will show significant deviations from the null hypothesis under consideration

The investigator is interested in testing for significant variation between groups as far as the averages or means are concerned He would also like to avoid incorporating into the experiment other causes for "significance" such as lack of homoscedasticity (lack of uniform variability in the scale) or lack of a normal (or Gaussian) distribution Therefore, before running an analysis of variance the distribution of data should be examined to see what type of distribution is present If it is not normal then a normalizing transformation must be considered

There is no particular reason why one scale should be used in preference to another in reporting the results of a biochemical analysis As a matter of fact the spectrophotometer has two scales to read One is called transmission and the other density The one is arithmetic and the other logarithmic Here then is a transformation built into the very dial from which the results are read Therefore, any scale can be used that satisfies the requirements of the statistical model

Since most of the analyses used in parametric statistics (statistics that depend upon a specific distribution) assume a "normal" distribution this distribution will be analyzed in some detail

Any distribution function $Y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$ of this form is called a normal probability density function or p d f with mean, μ and variance, σ^2 The title, "probability density function," implies that the area beneath the curve of this particular function is equal to one (1) and that since it is a continuous distribution, a precise probability statement can be made from a segment of the total area Because of the extensive experience with this particular model much analysis has been made to find out what random samples can be made to approximate the normal curve One powerful example of this is derived from the central limit theorem which says that a distribution of mean values tends toward normality as the sample size increases

When making deductions of a general character from samples drawn at random from an unknown population the researcher often needs to test whether a sample of values is likely to have been drawn from a Gaussian parent population In devising tests of normality two classes of procedures have been followed—(1) a normal curve is fitted to the sample data and the χ^2 test for goodness of fit is applied (2) Certain ratios of the moments of the sample are calculated and the significance of their departure from the expected value for a normal distribution is examined

First consider procedure "one" Although it has advantages it is somewhat insensitive to real departures from normality This is due to two characteristics of this test First, it takes into account *only* the magnitude of differences between the observed and expected group frequencies and not their sign or arrangement and, second, because of the necessity of grouping, the effect of small frequencies in the tail of the distribution is eliminated and errors due to grouping of data result *In the following,* the concern is with procedure "two" and the outcome of the specific normalization procedure used in this example

The statistical method for testing the assumption of normality using ratios of moments will be presented before the discussion of the actual numerical results This mathematical approach is based on the moments of the normal distribution First, let $X_1, \cdots, X_N$ denote the random sample from a parent population From this random sample a test is made to determine if the sampling comes from a normal distribution The first four moments about the mean of any population distribution are defined as follows

$$\mu_1 = \sum_{i=1}^{N} x_i/N$$

$$\mu_2 = \sum_{i=1}^{N} x_i^2/N$$

$$\mu_3 = \sum_{i=1}^{N} x_i^3/N$$

$$\mu_4 = \sum_{i=1}^{N} x_i^4/N$$

in which the deviations are $x = X - \overline{X}$ and where the mean, $\overline{X} = \sum_{i=1}^{N} X_i/N$ N here denotes the size of the random sample Now μ_1 is merely the mean of the deviations from the mean or as is sometimes called, the first moment about the mean μ_2 is the mean of the deviations squared from the mean or the variance μ_3 is the third moment about the mean and μ_4 is the fourth moment about the mean These moments distinguish certain types of distributions by certain ratios that can be formed from these moments When these ratios approach specific values, the random sample is known to approximate a sampling distribution that is normal or Gaussian

These ratios of the moments are called betas and are defined as follows

$$\beta_1 = \mu^2_3/\mu^3_2$$
$$\beta_2 = \mu_4/\mu^2_2$$

The two ratios, $\sqrt{\beta_1}$ and β_2 serve as a means of defining departure from normality These ratios are called sample statistics since they are derived from the random sample If the parent population is normal the sampling distributions of these two sample statistics tend toward normality as the sample size increases Their means are zero (0) and three (3) and their sampling standard deviations are *approximately* $\sqrt{6/n}$ *and* $\sqrt{24/n}$ *respectively Thus any significant* deviation from $\sqrt{\beta_1} = 0$ and $\beta_2 = 3$ indicates that the data does not have a normal distribution Furthermore, $\sqrt{\beta_1}$ is a measure of the skewness of the sampling distribution and β_2 is a measure of the kurtosis If $\sqrt{\beta_1}$ is greater than 0, the distribution has longer tails on the right (positively skewed) and if β_1 is less than 0, it has longer tails on the left (negatively skewed) β_2 indicates the general form of concentration around the mean If β_2 is less than 3, the distribution appears to be less heavily concentrated about the mean and if β_2 is greater than 3, it is more heavily concentrated about the mean

In order to evaluate each variable for its normality of distribution and then to transform it and finally to repeat the above evaluations with perhaps several repetitions requires long and laborious computations Therefore, without the assistance of automatic data processing one might be quickly discouraged, especially with 10 variables and 528 samples

Since no library program was available to test the assumption of normality one had to be written First the steps of such a program were outlined in a block diagram Then a series of FORTRAN or formula translation statements were written which was translated by the computer into a machine language program for computing $\sqrt{\beta_1}$ and β_2

The sampling distributions of $\sqrt{\beta_1}$ and β_2 have been determined by P Williams and E S Pearson, respectively From these determinations it has been found that if the sample is large, $N \geq 100$, departures from normality may be tested by treating $\sqrt{\beta_1}/\sqrt{6/N}$ and $(\beta_2 - 3)/\sqrt{24/N}$ as normally distributed values Hereafter these two normally distributed values will be denoted by z_1 and z_2, respectively

Figure 2 represents the flow chart to evaluate the normality of each of the ten variables under consideration The Fortran

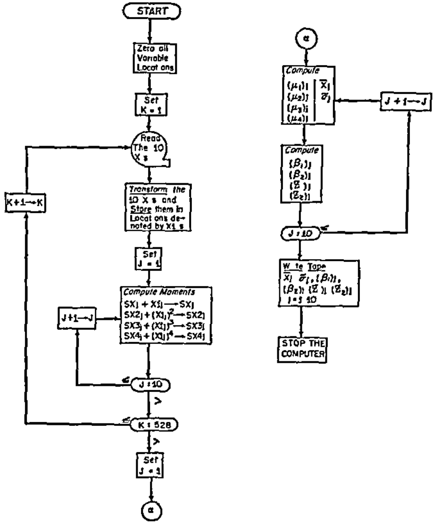


Figure 2 This is the flow chart of a program designed to calculate sample statistics which can be used to evaluate the frequency distribution of a set of data The actual FORTRAN statements are shown in Figure 3

→

Figure 3 These are the formula translation (FORTRAN) statements that were used to carry out the various steps of the analysis shown in Figure 2 The results of such an analysis program applied to a series of transformed criterion variables are shown in Table I

```
C  BULGREN-MASON—544-0004
   DIMENSIONSX(15),SX2(15),SX3(15),SX4(15),CSX(15),CSX2
    (15),1CSX3(15),  CSX4(15),TMEAN(15),TVAR(15),TMU3(15),
    2TMU4(15),BETA1(15),BETA2(15),  TSIG(15),Z1(15),Z2(15),
    3X1(15),X2(15),X(15)
   DO3K=1,528
   READINPUTTAPE14,1,(X(I),I=1,10)
 1 FORMAT(F4 0,2F3 0,7F4 0)
   X1(1)=LOGF(X(1)+200 )
   X1(2)=LOGF(X(2))/3
   X1(3)=LOGF(X(3)- 75)/2
   X1(4)=LOGF(X(4)- 75)/2
   X1(5)=LOGF(X(5)+100 )/2
   X1(6)=X(6)
   X1(7)=LOGF(X(7)+7 )/2
   X1(8)=LOGF(X(8)+700 )/2
   X1(9)=LOGF(X(9)+100 )/2
   X1(10)=LOGF(X(10)+625 )/2
   DO4J=1,10
    SX(J)=SX(J)+X1(J)
   SX2(J)=SX2(J)+X1(J)**2
   SX3(J)=SX3(J)+X1(J)**3
 4 SX4(J)=SX4(J)+X1(J)**4
 3 CONTINUE
   P=528
14 DO6J=1,10
   TMEAN(J)=SX(J)/P
   TVAR(J)=((P*SX2(J)-SX(J)*SX(J))/(P*P))
   TMU3(J)=(P*P*SX3(J)-3 *P*SX(J)*SX2(J)+2 *SX(J)*
    1SX(J)*SX(J))/(P**3)
   TMU4(J)=(P*P*P*SX4(J)-4 *P*P*SX3(J)*SX(J)+6 *P*
    1SX2(J)*SX(J)*SX(J)-3 *(SX(J)**4))/(P**4)
   BETA1(J)=(TMU3(J)**2)/(TVAR(J)**3)
   BETA2(J)=TMU4(J)/(TVAR(J)**2)
   TSIG(J)=TVAR(J)** 5
   BETA1(J)=BETA1(J)** 5
   Z1(J)=BETA1(J)/((6 /P)** 5)
 6 Z2(J)=(BETA2(J)-3 )/((24 /P)** 5)
   DO12J1=1,10
   WRITEOUTPUTTAPE3,21
12 WRITEOUTPUTTAPE3,22,J1,TMEAN(J1),TSIG(J1),BETA1(J1),
    1BETA2(J1),Z1(J1),Z2(J1)
11 CONTINUE
21 FORMAT(62H   VAR      MEAN      SIG       BETA1
    1BETA2      Z1      Z2)
22 FORMAT(I2,5H      ,6F10 5)
   END
```

seen in Figure 3 was written for the I B M 7070 Since Fortran is compatible with a large number of computers, this program could be used by any interested reader

FORTRAN is a system of processing statements which are written in a language which is understandable to man and ending *up with machine language statements which will run the computer* This system, which is stored in the computer, takes each FORTRAN statement and translates it into a series of machine language instructions Such a translation was carried out to produce the machine language program needed to carry out the flow *chart shown in Figure 2*

With this program it was possible to quickly evaluate all variables to be used If a transformation was required for normalization some suggestion was provided as to whether skewness or kurtosis (that is asymmetry, or abnormally peaked or flat configuration) *was the problem The transformation used was almost always of the* form $X' = \frac{\log (x + c)}{k}$ where c and k are constants The actual values of the constants c and k were determined by trial and repetitive evaluation of $\sqrt{\beta_1}$ and β_2 until a normal distribution *was approximated*

A very important point needs to be made at this stage regarding the use of transformations and analysis of variance It is possible to transform all significance out of any set of experimental data The supposition on which the null hypothesis is based is that the *experimental conditions cause no difference from the control conditions* in the criterion variable

If one were to use the experimental data and transform it so as to have a normal distribution, then if there were any real differences between groups it would be lost Therefore, transformations of all data for analysis of variance must be made on the basis of standard control data However, what is control in *one direction may be experimental in another direction of an* analysis of variance design This would appear to be an insoluble *problem and a serious objection to the analysis of variance method* unless the design is such that a very well defined standard control set of data can be selected for each analysis

The importance of a transformation to a normal distribution to obtain additivity This term refers to the property that when two objects, containing equal amounts of an attribute, are combined the third object is equal to the sum For example, the distance from two to four inches is exactly the same as the distance from 122 to 124 inches However, the area difference between a two and a four inch cube is twelve cubic inches, whereas the difference between a 122 inch and a 124 inch cube is 90,776 cubic inches This illustrates a scale in which additivity does not exist because the two differences added equal 90,788 and not 4

Table I illustrates the transformation used for each variable Associated with each variable is the transformation and the corresponding $\sqrt{\beta_1}$ and β_2 values along with the z_1 and z_2 which are the standardized normal scores described earlier Both the initial and transformed scale sample statistics are shown This shows the difference that the transformations made and how with the arithmetic values the assumption of normality may be far from valid

In the analysis of variance example under discussion the selected transformation was formulated around data collected during another (control) experiment in which normal mongrel dogs were eating the same type of diet and drinking water *ad libitum* and without the administration of any toxic agents

With the new scale the analysis of variance was ready to run Since an analysis of variance is a library program, all that is required to run a particular analysis is a few additional control instructions These instructions specify the input and output of the data, or the format as it is called, and the number of samples for each group under consideration These control cards adjust the library programs to the dimensions of the experiment under consideration by indicating the location of data to be read into the computer With the criterion variable being specified by the control instructions, all the other variables stored in the data cards or on magnetic tape will be ignored

We will now consider for purposes of illustration the specific computer center which contains our library routine The State University of Iowa Computer Center is at present chiefly a closed shop That is one in which special operators are required to operate the equipment This specific electronic digital computer is run by means of a monitor system with magnetic tapes as the

TABLE I

Sample Statistics

Variable	*Initial*				*Transformation*	*Normalized*			
	$\sqrt{Beta\ 1}$	*Beta 2*	*Z1*	*Z2*		$\sqrt{Beta\ 1}$	*Beta 2*	*Z1*	*Z2*
Volume	1 06	4 59	9 96	7 46	X' = Log (X + 200)	0 14	3 24	1 26	1 15
GOT	14 26	268 73	133 76	1246 38	X' = [Log (X)]/3	0 13	7 70	1 18	22 09
B-G	5 79	52 27	54 32	231 08	X' = [Log (X − 75)]/2	0 35	3 84	3 29	3 93
LAP	7 12	67 99	66 76	304 82	X' = [Log (X − 75)]/2	0 99	4 15	9 34	5 37
Creatinine	1 86	8 29	17 48	24 80	X' = [Log (X + 100)]/2	0 25	3 19	2 30	0 92
Sodium	0 01	2 44	0 12	−2 64	X' = X	0 01	2 44	0 12	−2 64
Potassium	1 79	6 92	16 76	18 39	X' = [Log (X + 7)]/2	0 21	2 97	2 01	−0 09
Chloride	0 37	2 91	3 52	−0 44	X' = [Log (X + 700)]/2	0 49	1 84	4 63	−5 43
Phosphorus	1 78	7 69	16 66	21 99	X' = [Log (X + 100)]/2	0 11	3 91	1 03	4 29
Total Solute	0 67	4 84	6 29	8 63	X' = [Log (X + 625)]/2	0 28	1 73	2 62	−5 93

only input-output The monitor is an automatic means of telling the computer to run in a specific and economic manner so that many different projects can be processed in one run A smaller computer is used for all the auxiliary functions associated with handling the data such as transferring data from punched cards to magnetic tape, printing the contents of the magnetic tape and for some of the preliminary sorting of the data in order that input time for the larger computer will be as short as possible

All that the user must do is to provide the center with the information as to the library program desired, the control statements or cards, and the appropriate set of data cards for the specific library program Twice daily the output is made available in the form of printed sheets of results If more rapid access to the results is needed, priority may be achieved for the investigator by contacting or actually working with the operator The results come from the small scale computer (I B M 1401) by way of the printer attached to it The printer yields tables covering sheets approximately twice as large as typing paper at a rate that will frighten the most avid seeker of results A computer can produce enough results in an hour to keep the research team busy for another six months Many people only use the computer for a few minutes at a time and many use it only once a month or less unless they are analyzing data from some ongoing experiment and have large volumes of data

The standard library program that was used for this particular analysis was the Type VI analysis of variance Only the analysis of LAP will be presented The output from this particular analysis was printed in the form of a table with the sums and means across for row column slices, row slices, column slices for group 1 dogs and then for group 2 dogs and then row column sums and means, column sums and means, row sums and means, slice sums and means and finally a summary table showing degrees of freedom, sum of squares, mean square and F ratios Table II illustrates the preceding by showing the printed output of the LAP analysis of variance from the computer The investigator may look for any "wild" values in these tables and if he finds some wild values he can then check back in the original data for errors in recording, such as misplaced decimal points The standard table of F-ratios

TABLE II

Type VI Anova Criterion = [Log (Lap − 6 5)]/2

Group C_1 High feed							
Row—Column—Slice Sums							
	6 08	5 10	5 74	6 43	7 05	6 53	8 10
	7 70	6 97	6 92	6 60	6 97	5 89	6 75
	5 90	7 11	7 45	6 59	6 33	6 86	6 89
	6 46	6 37	5 98	5 98	6 26	6 09	7 12
	7 14	6 32	6 84	6 55	9 98	7 28	7 49
	8 05	6 38	6 79	6 36	6 63	6 10	4 98
	5 14	5 20	8 15	5 96	5 82	5 26	
Row—Column—Slice Means							
	1 01	0 85	0 96	1 07	1 18	1 09	1 35
	1 28	1 16	1 15	1 10	1 16	0 98	1 13
	0 98	1 18	1 24	1 10	1 06	1 14	1 15
	1 08	1 06	1 00	1 00	1 04	1 01	1 19
	1 19	1 05	1 14	1 09	1 66	1 21	1 25
	1 34	1 06	1 13	1 06	1 11	1 02	0 83
	0 86	0 87	1 36	0 99	0 97	0 88	
Marginal Statistics							
Row—Slice Sums							
	80 19	78 58	85 09	72 77			
Row—Slice Means							
	1 11	1 09	1 18	1 01			
Column—Slice Sums							
	24 33	24 90	24 09	27 29	27 74	24 41	26 41
	26 31	32 00	26 62	26 28	26 26		
Column—Slice Means							
	1 01	1 04	1 00	1 14	1 16	1 02	1 10
	1 10	1 33	1 11	1 09	1 09		
Group C_2							
Row—Column—Slice Sums							
	5 93	5 49	4 77	4 43	4 72	4 55	4 54
	4 49	4 83	5 12	4 49	4 64	4 18	5 31
	5 07	4 45	4 33	4 01	3 74	3 85	4 86
	4 18	4 12	3 93	4 20	3 76	3 38	4 17
	4 59	4 44	4 62	4 21	4 66	4 07	5 18
	3 77	4 40	4 54	3 22	3 83	4 51	4 51
	4 64	3 19	5 61	4 69	3 96	4 08	
Row—Column—Slice Means					−2		
	1 19	1 10	0 95	0 89	0 94	0 91	0 91
	0 90	0 97	1 02	0 90	0 93	0 84	1 06
	1 01	0 89	0 87	0 80	0 75	0 77	0 97
	0 84	0 82	0 79	0 84	0 75	0 68	0 83
	0 92	0 89	0 92	0 84	0 93	0 81	1 04
	0 75	0 88	0 91	0 64	0.77	0 90	0 90
	0 93	0 64	1 12	0 94	0 79	0 82	

TABLE II (*Cont'd*)

Marginal Statistics							
Row—Slice Sums							
	58 00	52 04	51 03	51 18			
Row—Slice Means							
	0 97	0 87	0 85	0 85			
Column—Slice Sums							
	18 71	19 10	16 43	16 87	18 14	17 52	17 54
	15 74	19 97	18 06	17 75	16 41		
Column—Slice Means							
	0 94	0 95	0 82	0 84	0 91	0 88	0 88
	0 79	1 00	0 90	0 89	0 82		
Marginal Statistics R1 C1, R1 C2 R 4-C12							
Row—Column Sums							
	12 01	10 59	10 51	10 86	11 77	11 08	12 64
	12 19	11 80	12 05	11 10	11 61	10 07	12 06
	10 97	11 56	11 77	10 61	10 07	10 71	11 76
	10 64	10 49	9 90	10 18	10 01	9 47	11 28
	11 73	10 76	11 45	10 76	14 64	11 35	12 66
	11 82	10 78	11 33	9 58	10 46	10 61	9 49
	9 78	8 39	13 77	10 65	9 78	9 34	
Row—Column Means							
	1 09	0 96	0 96	0 99	1 07	1 01	1 15
	1 11	1 07	1 10	1 01	1 06	0 92	1 10
	1 00	1 05	1 07	0 96	0 92	0 97	1 07
	0 97	0 95	0 90	0 93	0 91	0 86	1 03
	1 07	0 98	1 04	0 98	1 33	1 03	1 15
	1 07	0 98	1 03	0 87	0 95	0 96	0 86
	0 89	0 76	1 25	0 97	0 89	0 85	
Marginal Statistics, Cube							
Column Sums							
	43 04	43 99	40 52	44 16	45 88	41 94	43 95
	42 05	51 97	44 68	44 03	42 67		
Column Means	0 98	1 00	0 92	1 00	1 04	0 95	1 00
	0 96	1 18	1 02	1 00	0 97		
Row Sums	138 19	130 62	136 12	123 94			
Row Means							
	1 05	0 99	1 03	0 94			
Slice Sums							
	316 63	212 25					
Slice Means							
	1 10	0 88					

provides the basis for determining the statistical results of prime interest

An F-ratio is a measurement of variation in a selected dimension of the experiment This ratio should approximate or approach unity if the factor designated does not cause any variation in the criterion variable An F-ratio is the ratio of two χ^2s, each divided by their respective degrees of freedom In an analysis of variance design Chi square (χ^2) is the sum of squared deviations from the mean divided by the vaianrce The variance is the mean square deviation of the criterion variable from the mean of the criterion variable It should be evident from these definitions why the test is called an analysis of variance

In discussing the results of the analysis of variance first consider the summary table The degrees of freedom indicates the number of opportunities within the designated dimension of design that the criterion variable may actually vary For example, there were eleven subjects so degrees of freedom (d f) for subjects is 10 That is after the mean is known and ten of the values for subjects are known then the last subject is also known There are, therefore, only n-1 degrees of freedom As another example, AB interaction has 11 x 3 degrees of freedom since there are twelve A columns and four B rows This is evident from the fact that $12 - 1 = 11$ and $4 - 1 = 3$ The d f are needed to look up in a standard F-table the inequalities required for a chosen level of significance as will be mentioned below

The sum of squares (ss) is an abbreviation for the sum of the squares of deviations of individual observations in a given dimension from the over-all mean The total variation of a given value from the over all mean is made up of variation due to the simple effects of A, B and C plus all the interaction effects Therefore, if one adds up all the separate ss it will equal the total ss in most analysis of variance tables

This is not true here because a number of separate error terms have been included for more reliable testing of the various interaction effects For example, error (w) has been subdivided into three components each of which is used to test a different simple or interaction effect Error (w) denotes the error within the groups The mean squares or (ms) are defined as ss divided by d f These are used to formulate the F-ratios of the last column Ss divided by the population variance is a Chi square and an F-ratio is defined as a ratio of two Chi squares over their respective degrees of freedom Since one of the basic assumptions of a Type VI design is that the population variances are equal for each treatment the population variances cancel out in their F-ratio

The F-ratio is a measurement of the relative variation between components as compared with the variation within components For example, to test the significance of the sum of squared variation from the mean of the 2 C columns, error B is used and hence we obtain 6 054/0 217 = 27 9

The F-ratios must then be compared with a standard Normally if the null hypothesis is true (that is if variation between components of the design show no significant difference from the variation within those components) then the F-ratio should have a normal distribution about unity Actually the F-ratio for the C dimension is 27 9 as seen from Table II According to the standard table for F-values, anything greater than 22 86 (for d f of 1 and 9) would only happen 1 per cent of the time, or once in a thousand such experiments, if the null hypothesis was true

From Table II the F-ratio for A is equally significant At this level of probability ($P < 001$) none of the interactions AB, AC or BC are significant As seen from Table II, we are left with C and A as highly significant from a probability aspect

It is now necessary to review what the actual relationship of mean levels is in these dimensions of the experiment and what was designed originally so that some conclusion may be drawn relative to the physiologic meaning, if any, of these findings

Since the data were transformed the means must be examined in the transformed state It is of importance to note what the transformation did to the data as compared to the original scale Some transformations will invert the scale In this instance the scale is still in the same direction and the lower concentrations of LAP were reported by lower means in the transformed data used in the analysis of variance See Figure 4 for the relationship of means

Now let us see if we can translate the analysis into simple words, that is with respect to the original conditions and all the steps taken This is where some judgment is required as to how much generalization is justified Certainly the results do not indicate what would happen in man if he were given a foreign protein

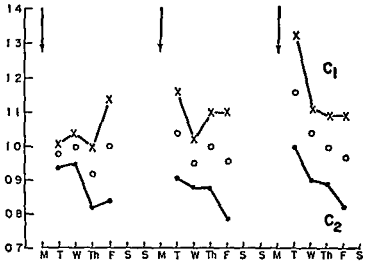


Figure 4 Shows the mean values for urine leucine amino peptidase (LAP) plotted in arbitrary units on the ordinate and days of the week on the abscissa Times of intravenous egg albumin injections are indicated by the Monday arrows Group C_1 dogs (6 in number) were fed more food on days Tuesday through Friday Group C_2 dogs (5 in number) were fed the same restricted diet on all days that C_1 dogs received on weekends The means for all dogs are shown by *circles* C_1, *crosses and* C_2, *dots*

or an altered food intake It can probably be said with considerable assurance that LAP excretion in the urine of dogs is reduced by a reduction in food intake and that regardless of whether dogs are eating a nearly normal ration or a greatly reduced ration, injection of egg albumin will cause an increase in the excretion of LAP

If LAP excretion is influenced by diet and by injection of egg white intravenously, is the latter effect to be interpreted as a toxic effect or a dietary effect? Simultaneous analysis of GOT (glutamic oxaloacetic transaminase) and B G (beta glucuronidase) revealed no significant effects of either diet or egg albumin injection In earlier experiments where neomycin was injected there was a rise of all three enzymes In circumstances where no nephrotoxin is given there is very poor correlation between the excretion of LAP and the other two enzymes From these observations the authors believe that the rise in LAP excretion after intravenous egg albumin in doses of 1 ml/kg and after the animals are sensitized, is not a toxic effect It is more analogous to the effects of diet and suggests that perhaps the important variable in the diet is the amount of protein Thus another hypothesis is made and new experiments will be required to test them

However, this is not the end of data processing in this illustrative experiment If urinary LAP was lower with restricted food intake, was there a change in some concomitant moiety in the urine that tends to be excreted with LAP? A multiple regression analysis was run by means of a standard library program to find out what significant relationships, if any, could be found LAP was used as the dependent variable

Table I shows the transformations developed with the illustrative experiment It is to be noted that for sodium no transformation was used in multiple regression analysis because the data was close enough to Gaussian as it was recorded in the data book Text books and papers written with statistical treatment of data often have the statement, "It is assumed that the data have a normal distribution" This statement is often erroneously interpreted to mean that the assumption can be made without testing it Just because a variable has a Gaussian (or normal) distribution in one set of data does not mean that it will be Gaussian in another experiment For example, Figure 5 shows a set of sodium data in which the mode was close to zero The transformation required was $X' = \frac{\log (x + 10)}{2}$ where X' is the f (x) and x is sodium When this same transformation was applied to the egg injection-experimental data under discussion the resultant distribution was almost as skewed but in the opposite direction

It is easier to decide what to do about scales in multiple regression or correlation analysis because all data may be used for determining the transformation required In order to obtain more of a feeling for the need of transformations let us consider the rationale behind a regression analysis A regression analysis is a method of fitting the data to the function of two or more variables

A common and general method used in estimating a population regression curve from sample data is the *method of least squares* The curve to be estimated is "assumed" to be a straight line There are several ways to test this assumption of linearity, but

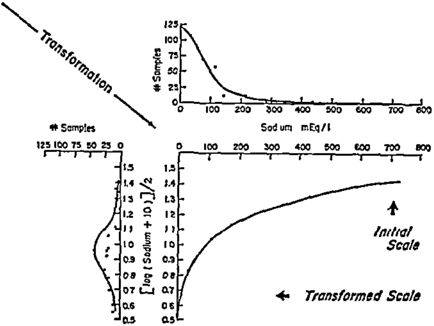


Figure 5 Shows a transformation from a scale of milliequivalents per liter to a new scale chosen because it would provide a closer approximation to a Gaussian frequency distribution

they will not be discussed here Briefly then, the sample regression curve of y (dependent variable) on x (independent variable) of given degree is the curve among all those possible of that degree that minimizes the sum of squares of vertical (y) deviations of the observed points from the curve In the linear case, the degree of the curve is one

By means of a multiple linear regression it is possible to determine whether several variables, x_1, , x_k, differ in their effect on a dependent variable, y Data gathered for regression analysis may be classified into either a type 1 problem or a type 2 problem The type 1 problem occurs when the effect on y may be measured for specific selected values of the variables x_1, , x_k, chosen by the experimenter The type 2 problem occurs when the points (x_1, , x_k, y) are chosen at random from a (k + 1)—dimensional normal distribution of points in the (x_1, , x_k, y) plane In either case the y value observed is considered to be a random observation from a normal population of possible y's for each particular x used These y populations are assumed to have a common standard deviation The assumption of normality is necessary for a significance test and confidence intervals in regression analysis Therefore, it is necessary to fulfill the basic underlying assumption of normality for the dependent variable in the type 2 problem

The multiple regression analysis program was written to do the analysis as described by Ezekiel and Fox This program is also a standard computer library program The output from this program was first a table of simple correlations between all variables, a multiple regression coefficient, the standard error, and the associated F-ratio Finally a table of partial correlation coefficients, partial regression coefficients and the associated partial F-ratios is given

With respect to the specific illustration experiment, the first run with 6 variables revealed that potassium and total solute concentration were important predictors of LAP while sodium, phosphorus and chloride concentrations were not It was observed, however, that if a nonparametric term was added to account for the effects of injection on LAP concentration, the multiple-R increased from 188 to 294 and the corresponding multiple-R F-ratio increased from 3 83 to 8 23 This was accomplished by assigning a number 3 to each sample collected on Tuesday, 2 for Wednesday, 1 for Thursday and 1 for Friday This is a means of weighting the days in relation to the Monday injections of egg albumin This new variable was called "injection factor "

A third regression analysis was then run with potassium, total solute (ΔT) and injection factor as the independent variables This analysis will be denoted as Analysis I The multiple-R fell from 294 to 277 and the corresponding F-ratio rose from 8 23 to 14 49 with the smaller number of variables selected

The regression coefficient squared is called the *coefficient of determination* Thus $(294)^2 = 086$ or 8 6 per cent of the observed variation in LAP could be projected from the independent variables when all seven were used and only 7 7 per cent when the data from potassium, freezing-point-depression and injection factor were used alone The 7 7 per cent predictability of LAP is more significant, however, because it is accomplished with a smaller number of terms

The fourth and last regression analysis run for this experiment was the same as Analysis I except that two more independent variables were added The first variable added to the regression equation was creatinine, since it was known from previous analyses that it should increase the multiple-R and the corresponding F-ratio The second variable added to this equation was "GOT" GOT should be lost from cells when they are damaged and to the extent that LAP loss is a toxic or cell damage effect of the egg injection there should be correlation This analysis will be denoted as Analysis II With these two independent variables added, the multiple-R increased from 277 to 371 and the corresponding F-ratio increased from 14 49 to 16 69 The coefficient of determination now is 0 137 Almost 14 per cent of LAP variation in this experiment can be related to variation in GOT, creatinine, potassium, total solute and "injection factor"

For illustrative purposes, the print-outs obtained from the computer are reproduced in Table III and Table IV These two regression analyses are from Analysis I and Analysis II respectively The "echo check," found at the top of Tables III and IV, is a listing of the first set of variables from the experiment These values have all been transformed by the obtained normalizing factor in order to fulfill the assumption of normality The transformations used in both analyses are $LAP = [\log (X - 75)]/2$, $Cr = [\log (X + 100)]/2$, $GOT = [\log X]/3$, $K = [\log (X + 7)]/2$, and $\Delta T = [\log (X + 625)]/2$ The X denotes the arithmetic values of each random variable "Injection factor" is a term representing the time-related introduction of the foreign protein In a sense the analysis is one of autocorrelation It is somewhat like the evoked response correlogram type of study used in electroencephalography only instead of looking for brain wave correlation dependent upon time the present analysis searches out relationship of LAP excretion to time of introduction of the foreign protein and to any time-related excretion of other moieties

The mean values for the analysis of each variable are listed in the print-outs along with "population standard deviation," which is an estimate, and "standard deviation" which is the value observed considering the data as a universe and not as a sample Again, these are averages and standard deviations calculated from the transformed values

This still perhaps may be a stumbling block to some How can you add all values and divide by the total number if the values have been changed from their original form by taking their logarithms? Adding the logarithms of numbers is the same as multiplication of the original values The answer to this is that the original scale is no longer of interest

The transformation was for the purpose of providing a new scale in which the distribution of sample data would be Gaussian Having changed to this scale, the analysis is carried out with the new values and the results will be interpreted in terms of the new scale The new values can be added and a mean determined, and the mean will be applicable only to the new scale

If one wished for some reason to see where this mean would fall on the old scale it would be necessary to take the antilog and go through the other manipulations indicated by the transformation used for that variable This would then be a sort of geometric mean of the original data The calculation of standard deviation for this "geometric mean" would require adding and subtracting the standard deviation from the mean using the values in Tables III and IV and then taking the inverse of the transformation for all three values

To actually extract and interpret values from the regression analyses carried out with a transformed scale and to try to determine their value in the original non-Gaussian scale is dangerous, if not impossible, and is without purpose One scale is as good as another The investigator should not be so emotionally attached to one scale that all thinking must be done in that scale Actually the statistical analyses lead to ratios or pure numbers which are without any scale except for distribution frequency

The correlation matrix is a table of simple correlations with every variable correlated with every other variable The highest correlation coefficient observed in the experiment is between total solute and potassium (75) In Analysis I, the best simple correlation with LAP appears to be the "injection factor" term, with a correlation coefficient of 0 21 In Analysis II there is a slight

TABLE III

Regression Analysis I

No of Observations = 528 No of Variables = 4

Echo Check

0 9039	0 9375	1 5692	1 0000

Variable	*Mean*	*Pop S D*	*St Dev*
LAP	1 02	0 24	0 24
Potassium	0 82	0 15	0 15
Total Solute	1 57	0 05	0 05
Inj Factor	1 25	0 60	0 60

Correlation Matrix	*LAP*	*Potassium*	*Total Solute*	*Inj Factor*
LAP	1 00	0 02	0 12	0 21
Potassium	0 02	1 00	0 75	0 00
Total Solute	0 12	0 75	1 00	−0 08
Inj Factor	0 21	0 00	−0 08	1 00

	Mlt R	*Std Error*	*F*	*DF*	*Dep Var = LAP*
	27678	0 23014	14 49052	524	
	B	Beta	F	Part R	
Potassium	−0 30	−0 20	9 36	−0 13	
Total Solute	1 33	0 28	19 34	0 19	
Inj Factor	0 09	0 23	29 24	0 23	
A	−0 93				

negative correlation between creatinine and LAP which would suggest, but certainly does not prove, that LAP comes from kidney tubule cells in contrast to creatinine which comes from glomerular filtration There are few conclusions that can be drawn from simple correlations but the investigator will learn from experience and by study of such tables in repeated experiments that there are certain moieties that are usually correlated and this background of knowledge may occasionally allow him to pick out some unusual findings in a paricular experiment

In addition, if very high correlation coefficients are observed it suggests that these two variables are probably measuring the same basic factor The simple correlation matrix is used in the computer library program for the multiple regression analysis and, therefore, represents results of part of the analysis in progress

The final results are summarized in the line which lists the multiple regression coefficient, standard error of the multiple

TABLE IV

REGRESSION ANALYSIS II

No of Observations = 528 No of Variables = 6

Echo Check

0.4263 0.9039 1.3116 0.9375 1.5692 1.0000

Variable	*Mean*	*Pop SD*	*St Dev*
GOT	0.37	0.07	0.07
LAP	1.02	0.24	0.24
Creatinine	1.29	0.10	0.10
Potassium	0.82	0.15	0.15
Total Solute	1.57	0.05	0.05
Inj Factor	1.25	0.60	0.60

Correlation Matrix	*GOT*	*LAP*	*Creatinine*	*Potassium*	*Total Solute*	*Inj Factor*
GOT	1.00	0.24	0.24	0.34	0.34	0.08
LAP	0.24	1.00	−0.07	0.02	0.12	0.21
Creatinine	0.24	−0.07	1.00	0.65	0.57	0.00
Potassium	0.34	0.02	0.65	1.00	0.75	0.00
Total Solute	0.34	0.12	0.57	0.75	1.00	−0.08
Inj Factor	0.08	0.21	0.00	0.00	−0.08	1.00

	Mlt R	*Std Error*	*F*	*DF* *Dep Var = LAP*
	0.37128	0.22281	16.69	522
	B	Beta	F	Part R
GOT	0.75	0.23	26.67	0.22
Creatinine	−0.41	−0.17	10.43	−0.14
Potassium	−0.23	−0.15	4.73	−0.09
Total Solute	1.26	0.27	17.77	0.18
Inj Factor	0.08	0.21	26.02	0.22
A	−0.63			

regression coefficient, F-ratio for multiple-R, and degrees of freedom for the regression analysis (Tables II and IV). The multiple "R" is the correlation coefficient for the dependent variable (LAP) observed values compared with the values which would be predicted from the observed independent variables of creatinine, GOT, potassium, freezing point (ΔT), and injection factor for each urine sample using the derived regression formula.

In regression analysis, the *standard error* $S_{y/12 \ldots k}$ is a good estimate of the scatter of the population about the line or population regression plane. If the deviations from the plane are normally distributed, about 95 per cent of the points in the sample will lie within $\pm 1\ 96\ S_{y/12\quad k}$ of the plane These deviations are measured in the y (dependent variable) direction In Analysis II (Table IV), $S_{y/12\quad k} = 223$ whereas $S_y = S_{LAP} = 239$ If the deviations are approximately normally distributed, about 68 per cent or 359 of the 528 observed values of LAP should lie within 223 of the values predicted from the corresponding independent variable values by the regression equation It also appears in Analysis II that the predicted value for LAP has about the same standard error as the observed standard error for LAP determinations in this experiment

Only a restricted interpretation can be made from the model The variation observed has been determined by the way in which the experiment was conducted The correlation model applies here only in so far as future experiments are carried out in exactly the same way The experiment seems more one of multiple correlation than of multiple regression since no variables were fixed except the "injection factor" which is an independent variable An F-ratio of 16 69 can be evaluated for the multiple correlation model in the same way that it is used for the variance model The F-ratio associated with the multiple-R is used to test the hypothesis that the population multiple-R is zero Since F = 16 69 with 5 and 522 d f is larger than the critical value F_{05} (5,522) = 2 21, we reject the null hypothesis and conclude that the regression of LAP on GOT, creatinine, potassium, freezing point and injection factor does account for a significant amount of variation observed in LAP

Another manner of observing the amount of the variation in the dependent variable related to the chosen independent variables is by obtaining the coefficient of determination, which is equal to r^2 In the Analysis II, $(0\ 37)^2$ approximates 14 per cent Fourteen per cent of the LAP variation can be accounted for by the variation in creatinine, sodium, phosphorus and "injection factor" using the observed transformed values of these independent variables and the coefficients listed under "B" in Table IV For example, the equation would be LAP = 0 75 (GOT) − 0 41 (Cr) − 0 23 (K) + 1 26 (ΔT) + 0 08 (IF) − 0 63 This equation would be repeated 528 times with the transformed values inserted each time and the resultant, estimated LAP would account for about 14 per cent of the observed LAP

The "Beta" values are coefficients which would be used in the equation if each of the terms were divided by the standard deviation for that term

The F-values listed and the partial correlation coefficients have *about the same significance for individual terms as the multiple*-R and F-value listed for the over-all equation The difference here is that the degrees of freedom for the partial-R's are 1 and 522 The coefficients of determination can, however, be obtained by squaring each partial r Thus it appears that an injection factor partial r of 0 22 gives a D of $(0\ 22)^2$ or 4 8 per cent Therefore, almost one-third of the predictability of this equation in determining LAP *is dependent upon the term "injection factor"*

It still cannot be decided whether the LAP loss which occurs in relation to injection of egg albumin is a "toxic" effect or perhaps an effect of increased protein in the body similar perhaps to the dietary effect observed in the C *dimension of the analysis of* variance Even if GOT represented a toxic effect in the 5 per cent coefficient of determination, there is still 95 per cent of LAP variation that could be nutritional It is evident that the analytical problem here may well be autocorrelation Factor analysis might be a possible answer to this but has not yet been used for this experiment A statistical model that has been tried with this experiment is the covariance model

In covariance analysis the C_1 and C_2 groups were analyzed as high and low feeding groups with LAP as the dependent variable and with GOT, creatinine, potassium, freezing-point-depression and *injection factor (IF) as independent* variables We know that there will be a significant difference in the LAP means between C_1 and C_2 This was determined in the analysis of variance It is also known for this experiment that all of the above independent variables are significant predictors of LAP because of the results of regression analysis The question, however, is to find out how much nutrition influences the slope of LAP as predicted by the independent variables If GOT and injection factor are more related to the toxic effect they might not be influenced by the marked *difference in food intake of the two groups of dogs*, C_1 and C_2

GOT as a criterion variable in the Type VI analysis of variance showed no significant variation due to diet or injection of egg albumin From the regression analysis (Table IV) where the variables GOT and creatinine were added, the F-ratio associated with GOT is much more significant than the F-ratio associated with creatinine This is also seen from the two partial-"R"s

Two covariance analyses were run, one with LAP as dependent and potassium, GOT, total solute, creatinine, and IF as independent variables and the other with creatinine deleted Both results showed significant difference of regressions The results with creatinine included are shown in Table V

GOT and "injection factor" are not interchangeable terms in spite of having similar coefficients of determination in the multiple regression analysis The significance of the regression analysis comes from data collected from the C_1 group of animals The striking effect of food intake on LAP excretion influences the regression analysis by simple reduction in LAP excretion The results also suggest that there is minimal "toxic" or damage effect of injecting intravenous egg albumin and that perhaps the main stimulation of LAP excretion is "nutritional"

The generalizations that can be made from these analyses of one experiment must be limited Because of other experimental data and also because of the highly significant results observed in the C dimension of the Type VI analysis of variance and in the difference of means of the covariance analysis it is safe to predict that in future experiments in dogs nutrition will be an important condition to consider in the design when LAP is the criterion measure under study However, nothing can be said about the LAP excretion in man

The main thing to be learned from this example is the possibility of applying different statistical models to a set of data when the experiment is well designed The final interpretation is dependent upon a combined interpretation of all models plus what has been learned from other experiments, what is known about the actual organ system under study and information from the literature

The data processing is only in part looking for significant statistical differences The first objective is the stimulus to thinking

TABLE V

ANALYSIS OF COVARIANCE
2 GROUPS
5 INDEPENDENT VARIABLES

	GOT	*Creatinine*	*Potassium*	*Total Solute*	*Injection Factor*	*LAP*
Group Means, N = 288	0 38	1 27	0 82	1 57	1 25	1 11
Group Regression *Coefficients*	0 86	−0 40	−0 33	1 48	0 10	
Group Regression Sum of Squares					3 26	
Group Means, N = 240	0 37	1 31	0 83	1 57	1 25	0 91
Group Regression Coefficients	0 17	0 41	−0 16	0 31	0 05	
Group Regression Sum of Squares					0 61	
Over all Means, N = 528	0 37	1 29	0 82	1 57	1 25	1 02
Over-all Regression Coefficients	0 75	−0 41	−0 23	1 25	0 08	
Over-all Regression Sum of Squares					4 14	
Combined Regression Coefficients	0 65	−0 04	−0 27	0 89	0 08	
Combined Regression *Sum of Squares*					2 92	

	DF	*Sum of Squares*	*Mean Square*	*F-Ratio*
Over-all Regression	5	4 14		
Difference of *Positions*	1	4 11	4 10	98 0
Difference of Regressions	5	0 95	0 19	4 7
Combined Residual	516	20 86	0 04	
Total	527	30 06		

SUMMARY TABLE

Source of Variation	*DF*	*Sum of Squares*	*Mean Squares*	*F Ratios*
Dogs	10	8 003	0 800	
C	1	6 054	6 054	27 9
Error B	9	1 950	0 217	
Within dogs	517	26 093	0 050	
A	11	2 052	0 187	5 2
AC	11	0 840	0 076	2 1
B	3	0 923	0 308	1 7
BC	3	0 703	0 234	1 3
AB	33	2 605	0 079	2 8
ABC	33	2 377	0 072	2 6
Error (w)	423	16 592	0 039	
Error 1 (w)	99	3 585	0 036	
Error 2 (w)	27	4 766	0 177	
Error 3 (w)	297	8 242	0 028	
Total	527	34 097	0 065	

that comes from orderly assemblage and reduction of data and the variety of viewpoints The second objective is to consider the requirements of each model and to be aware of the distortion and erroneous results that may come from applying Gaussian statistics to non-normal data or data that lacks homoscedasticity No single experiment will be crucial Experience with numerous related experiments improves the ability of the investigator to select appropriate statistical models, parameters and scales and to generalize from the results

This type of analytical experimentation with the data would be impossible without computers However, if the investigator will maintain his perspective and if each trip to the computer leads to added insight and the stimulation of new questions then the authors have faith that these powerful techniques can be applied to clinical problems and that the rules and regulations will be found for assisting in diagnosis, prognostication and treatment evaluation with some of the rigor that exists in other disciplines The available analytical models have not been exhausted in this example Statistical and mathematical methods have required too much time by hand to have full application to very many of the complex medical problems and to be applied by physicians who know these problems Computers open the door to a much more thorough study of medical data while it is fresh in mind and in a way that can be fun as well as informative

LIBRARY
UDAIPUR

Chapter VII

INFORMATION ACCESS

THE physician with his patient and a specific question at a specific time and place is the focal point of information flow in medicine A physician may shift his emphasis to research or teaching but still his main concern is the patient He stores information about the patient in the record room He may attempt to assemble information from similar patients from other records There are many similarities between the record room problem and the medical library problem, but this chapter is directed primarily at the storage and recovery of pertinent medical literature

Information flow is required for and closely related to creativity Given average intelligence and sufficient drive, the best chance of being creative is to gather together, assimilate and retain as large a sample of the world's knowledge pertinent to a problem as can be accomplished in the span of one's attention This alone would justify the current activity in library methods and information science A second stimulus to the automation of our library systems is the exponential, inundating growth of literature

The requirements of a successful information storage system might be classified as (1) a basic understanding of what information is and a system for coding it, (2) an efficient, completely automatic storage system which is economical in cost and space and which provides a quick random access to its contents with facilities for duplication of any documents which are to be distributed, and (3) a computer program for discriminatory retrieval of stored material

Most physicians have so far noticed little change in their habits of library use They continue to subscribe to a few journals Some physicians are subscribing to Current Contents, a weekly report of the tables of contents in their original format, reproduced from more than 600 foreign and American journals and covering more than 125,000 articles by title It is possible to skim through these titles by the expenditure of an hour or less each week The information scientist will usually transfer those titles of interest to 4 x 6 inch reference cards and these are filed away in his personal filing cabinet which is rapidly climbing toward the office ceiling

As time permits, and demands require, segments of this file are taken to the library and the original articles skimmed or read in detail In a few instances reprints are requested by mail and further time is spent in filing the reprints

Indices such as the Index Medicus, Chemistry Abstracts and Biological Abstracts are occasionally consulted Additional information is transmitted by the convergence of physicians at the bedside of particular patients and perhaps after the presentations of certain papers at national meetings

During the year that this book has been in preparation there have appeared some key-word in-context indices which are the product of computer use and programs such as that prepared by H P Luhn One of the first of these was the *KWIC Index to Neurochemistry* which includes 2100 titles of papers permuted according to the key words plus some added descriptive identification into some 15,000 alphabetically arranged entries The first entry title contains the word *adrenaline* and the other key words are presented in sequence, but the word *adrenaline* is placed in the center of the line This same entry shows up elsewhere with *catechol* shifted to the center and is also found under *amines* and under *methylation*

At the end of each entry is an alphabetical numeric index which permits the user to find the reference in a separate section where they are grouped according to categories and then by alphabetical order of the senior author There is also an author index and an index for addresses by senior authors

The key words used come from the title or are added by those preparing the index These words are selected for their descriptive or informational content of the subject matter in the paper

This initial volume was prepared as an experiment The papers selected for inclusion were published between January 1960 and July 1961 This represents an interesting experiment from another standpoint in that it was a co-operative endeavor between a foundation in Amsterdam and two computer centers in New York The idea was conceived and made a reality in the short space of three months Other similar uses of this technique will be mentioned below, but first a note about the current intensity in research and development

The National Science Foundation has two current publications which serve to keep information scientists abreast of developments Scientific Information Notes is published every two months and contains information about current meetings, conferences, research and development, scientific information activities abroad and publication notes about new journals, indices, directories and other publications of potential interest to the scientific community

The second publication of the National Science Foundation of interest is published every six months and is titled Current Research and Development in Scientific Documentation (CRDSD) This is cumulative The last issue, Number X, includes information about 450 research projects being carried on in over 200 different organizations This represents an increase of approproximately 100 in the last six months This report describes studies of scientists' information needs, uses made of scientific information, communication problems, new methods of storage and retrieval, mechanical translation of languages, equipment design, studies of automated character recognition, speech analysis, and many related studies In addition there are presented brief statements reviewing the progress made in these projects and cross references to related studies A glossary of special terms has been added in the last edition

Another similar illustration of the magnitude of the current activity is a bibliography prepared by the American Institute of Biological Sciences, edited by P C Janaske, entitled Information Handling and Science Information which covers 1121 references between the year 1957 and 1961 This also has a permuted index in which each title is repeated in the appropriate alphabetical sequence for every key word in the titles of papers Brief abstracts are printed with each entry in the bibliography

The ultimate objective of all of this activity is to make it possible for an individual to obtain quickly a personalized index of the available articles of interest to him, along with a personal file of the original paper or its translation if the original was in a different tongue from the tongue of the user

L E Allen has suggested an approach to the cataloging of the entire body of written knowledge This has seemingly been a dream of mankind since the invention of writing and was attempted early in the development of civilization However, even the invention of the printing press and the publication of large encyclopedias failed to keep pace with the increase in knowledge and it is only *now with the invention of computers that this idea again seems feasible If it ever is attempted there will still remain the* "classified" information which, according to G Piel, already fills more than 100,000 file drawers in the city of Washington and at various *United States military and governmental installations*

More pertinent to the problems of medicine are the 250,000 papers a year which are currently cataloged in the National Library of Medicine The focal point for information in medicine is the physician as he stands in time and as he considers his role in society at that time He may have just received a phone call about a difficult patient with an unusual disease or complication He is suddenly extremely interested in obtaining all of the information available which will help him provide an optimum basis for his decisions and actions in the care of this patient

Later his role may change to that of an investigator, interested in retrieving all available information about similar patients so *that he may add his own observations and draw his own conclusions and perhaps contribute some new bit of information to the store of knowledge* This new information may arise because of his ability to gather together a larger sample of the world's knowledge related to this problem It may result from his new viewpoint because of the peculiar circumstances of his exposure, a new dimensionality in experience

He may be stimulated to undertake some laboratory experimentation related to this particular problem If so his informational requirements may shift and expend He may wish to retrieve information from a number of related fields, some of which may even be outside of medicine

When he finally reports the results of his research he must analyze his paper and summarize it He must gather together the total informational content and carefully choose words which will catch the eye of those informational scientists working in the same area This is his summary, but finally he must be clever enough to select words for the title of his paper which will include most of the information that is present in the summary

Most practicing physicians are concerned most of the time with information retrieval, but they must go through some of the same thought processes in retrieving information as the writers went through in coding the information in the titles of their papers Arnold Lazerow provided the authors an illustration of the problem as it pertains to diabetes literature The National Library of Medicine is classifying its one quarter million articles per year under some six or seven thousand subject heading entries Dr Izzo at the Center for Documentation and Communication Research at Western Reserve University in Cleveland has shown that by the use of 121 of these subject heading entries it is possible to retrieve nearly all of the diabetes related literature, but this still leaves 20,000 articles a year of which only about 2500 are actually of interest

A second level of retrieval is being developed which will require specially trained people to actually read through the papers and determine whether they contain experimental data, whether control groups were used, whether statistical analyses were applied and whether significant differences were found The articles will probably also be indexed as to their geographic scope of circulation, the type of audience for which they seem to have been prepared The study at present is attempting to determine what types of information about articles are necessary to identify them in ways which will be useful to practicing physicians who are interested in obtaining pertinent articles related to their problems

A third level of retrieval is under study at the University of Minnesota under Dr Lazerow and in this the objective is to develop methods of codifying papers of interest to the investigator In this type of retrieval selected papers will need to be abstracted with sufficient information that rather specific questions can be asked such as "Does growth hormone administration change the insulin content of the pancreas?" "Does growth hormone administration affect the granule content of the beta cells of the pancreas?"

It has been pointed out in connection with this depth of retrieval that the reduction in clerical labor in obtaining bibliographies might encourage some of the more imaginative investigators to undertake the writing of review articles and that such automation of information dissemination would lead to savings of time and money in research The average individual United States Public Health Service Grant of $25,000 a year has been used as an illustration by Lazerow If a six month delay occurs before an investigator discovers that a published article contains the answers for which he is spending his research money the loss due to duplication of effort could cost at least $10,000

The actual experience of working on a research project for several years only to find that all of this has previously been done has been a powerful stimulus to the development of better methods of information flow It has also been a powerful stimulus to try to combat the tendency of our government to "classify" the products of research

At present we retrieve information by title The words in the title are permuted and arranged in the pattern familiar to users of the Index Medicus The state of the art has been increased immeasurably during the last year by the appearance of B A S I C which is a key to the world's biological research published in *Biological Abstracts* B A S I C is a new unit This system, designed for use at the information scientist's own desk, is published every two weeks and is a permuted key word title listing similar to the Index on Neurochemistry mentioned above The scientist uses the key words which he selects according to his personal interests and, of course, these change with time It is still necessary when current papers of interest are found to then go to the nearest library or available *Biological Abstracts* to find the reference and then the original paper must be looked up in the current issue of the journal

Titles of papers are used for the preparation of B A S I C and each is entered on a punch card for computer processing Each basic word is ordered and turned to the center of the line and assembled in alphabetical order with a vergule (/) marking the end of the title The computer program carries out all of the arranging and prints out copy ready for photostatic reproduction and offset printing so that any required number of bound copies of the abstract can be prepared Words containing more than one idea are permuted For example, the word hemoglobin in a title would cause that title to show up under hemo and also under globin When a word has two components but only one idea it is hyphenated and treated in the program as a single word The availability of this bound key to the current biological research is a giant step in communication because it shifts the focus to the individual investigator in his own office and away from the local library with its increasing inadequacies

There are two obvious steps that should follow and which are under study S Zusman and R S Ledley have developed a biomedical information retrieval program called Tabledex which would permit a computer production of a personalized bibliography in bound form for use at the worker's desk and which would facilitate the more rapid retrieval of papers from a selected segment of the world's literature

The other development is illustrated by the Walnut System prepared for the Central Intelligence Agency which makes possible by compressed microimage storage and random retrieval, access to millions of documents from which copies could be automatically prepared and transmitted to the investigator in the form of windowed filing cards, the window containing a 3 x 5 cm photomicrograph of eight pages of printed material at 1/20 or up to 1/40 reduction Ultimately a system will undoubtedly be developed for immediate wire transmission of the image from storage to the user's desk viewbox and there will be no need to keep a separate file in one's office

During the last year the United States Department of Health, Education and Welfare, Public Health Service, National Institutes of Health, Division of Research Grants, has prepared a sample index of research grants that had been awarded in 1960 with the key word type of program This was prepared with modifying words or *descriptors* which were selected by their staff This required considerable additional time and, therefore, the 1961 grants were prepared in a Research Grants Index using only the titles of the grants

There were 7000 key word headings and phrases used to alphabetically arrange the 14,000 titles of current projects This catalog was then distributed to each of the grantees so that they could have at their desks a rapid access to information about the entire research effort being supported by this agency

It was then decided that although more professional time is required for its preparation a greater amount of more concise information could be obtained by using *descriptors* and these will be used in the 1962 index instead of the title words The descriptors are words taken either from title or from the information available about the research project by the professional personnel This then is the current state of the art and it is obviously in a great state of flux

The great advance in the next decade will undoubtedly come from the application of information theory to these problems Just as harmonic analysis involves distinguishing rhythms or redundancies of varying frequency in analysis of a curve so information theory seeks for the basic amount of information in such diverse signals as language, music, color patterns or complex smells In a sentence there is a certain predictability of the next letter given one, two, three or N number of preceding letters There is also a certain predictability of word sequence, and there are larger rhythms that cover paragraphs, chapters, books and even fields of endeavor according to the Zeitgeist

Information theory opens a Pandora's box Up until this time indices have been composed essentially of descriptive words taken somewhat passively from the titles of papers The levels of retrieval mentioned in connection with the diabetes study begin to raise some interesting questions about how much information there is in a word or a sentence Information theory tells us that these questions can be answered

According to the theory, information represents the number of bits (binary digits) required for predicting a letter or a word, etc The total information has also been referred to by the term "entropy" Actually man's search is for new patterns, new redundancy in knowledge which reduces the entropy

What would an encoder need to know about a subject in order to determine the informational content of a new paper? For example, if the first letter in the title were T the probability is greater than 1 in 26 that the next letter will be an H If the next letter turned out to be something other than H this would have higher informational content Given two letters, TH, the probability is high that the next letter would be an E Information theory has to do with recognition of the patterns which reduce the equivocation as to the next letter or the next word or the next idea

How many polar-option decisions are required to uncover all of the information in a title? How would an encoder begin to analyze a title, "The Clinical Significance of Serum Sodium Concentration"? The word of greatest convergence is *sodium Clinical* and then *serum* reduce the informational content tremendously first by eliminating almost the entire field of pure chemistry and secondly by eliminating any direct concern with cell fluid, bone sodium, urine sodium, or other potentially important clinical sodium patterns of information

The convergence patterns in the collection of words *clinical serum sodium* would require, however, that the encoder read the paper and determine how many of these patterns are actually present and to what degree That there are pattern implications of the term *clinical serum sodium concentration* is emphasized by the word *significance* in the title Some important informational patterns might be summarized in some additional groupings such as "The rule of electrical neutrality" or "The rule of osmotic equilibrium" or "The Gibbs-Donnan Effect" or simply "Tonicity" These various thought patterns converge in the term *sodium*, but they do not necessarily diverge automatically in the mind of the average reader because of the great variation in training, in viewpoint, in past and present experience, in the currently available mental maps that exist for each individual who sees these words

The encoder might, therefore, need to determine how much of each pattern was present in the article for some theoretical average reader Thus there are difficult problems in encoding the informational content relative to redundant information and even greater difficulty arises relative to judging the content of new information patterns or *surprisals* There is in addition the problem of the increasing information which arises as one's viewpoint changes and as the dimensions of one's experience and observation increase With the publication of a new paper containing really significant new information it is to be anticipated that a review of older papers on a subject will reveal additional information because of the change in dimensionality that has occurred

There is also the problem of evaluating information If a paper contains a new idea, a surprise, a few bits of information that could not have been predicted, is it because the author is dreaming without the hindrance of investigation or experimentation? Is the *surprisal* reproducible? Does the author present data? Is there an adequate control study? How significant is the conclusion reached from a statistical viewpoint? Is the statistical model used appropriate to the problem?

Obviously considerable empiricism must remain in the encoding judgments at least in the present state of the art Fortunately the empirical model of man's thinking is the ever present subject of our deliberations and there seems to be some safety in accepting certain "common sense" decisions until more objective and optimal encoding techniques are proven Today the title of a paper plus some added key words, selected by specialists in the subject in co-operation with the author, should make possible encoding which will be discriminative

It would appear that information science has a great growth potential, will provide many jobs requiring some shift in educational preparation from what has been required in related jobs There are undoubtedly many research projects where teams of scientists have come together from various disciplines that could benefit from the addition of a well trained information specialist This will be hard to accept since we all are information specialists, but as research becomes more complex we add chemists, physicists, electrical engineers, statisticians, systems specialists, programers, etc , etc , why not an information specialist? This is not just a gimmick to raise a secretary's salary They are highly trained and practically non-existent In fact they probably aren't defined, but that is one of the difficulties with progress

Chapter VIII

EDUCATION IN PROBLEM SOLVING

A CURRENT revolution in time-saving machines has some of the same implications for change that existed in the Industrial Revolution from labor-saving machines Machines can be used for storing and retrieving information and making routine decisions The initiation of any new program along this line requires that very human quality, creative thought

Time-saving machines can relieve man of his last remaining drudgery so that he may be really free to be *creative* and to perform those purely human activities which machines cannot provide

If creative thinking is to be a common commodity there will need to be a never-ending and intensive educational program at least for those members of society who are willing to participate in the current time saving revolution The development of automation is well established, accelerating in pace, irreversible and a barometer of an advanced, healthy, modern society The choice is not whether to develop computer techniques in our time The only choice is an individual one, namely, whether to participate

Evolution has always been difficult for some members of society but never before has the challenge and the opportunity been so great for the individual Today the individual can, through properly channeled hard work and education, reap the benefits of automation whether he is with or without status or riches

It is becoming increasingly apparent that man's most precious commodity is his own time With free time an individual may develop his information handling ability He may increasingly control his destiny by having control over his environment and himself Modern man seeks power to know and to make reliable decisions As long as he avoids the pitfall of subverting these drives to controlling his fellow men, today's citizen and society are increasingly fortunate

The challenge for physicians is to make the practice of medicine ever more scientific and at the same time more humanly attractive These two goals are more than compatible They are supplementary Practicing physicians must seek to attain as much of the logical, mathematical and statistical education as can be accomplished in their busy lifetimes They must introduce these quantitative and objective approaches with increasing emphasis in current medical education and they must also encourage the changes which are occurring in undergraduate education and which seem to be desirable in preparing students for a life of change and adaptability

The great artist must learn something about the medium in which he wishes to work The physician who wishes to really understand, and especially if he wishes to contribute to the understanding of his colleagues, must, like the artist, determine what medium is best suited to his problems and then must learn as much as is necessary about his chosen tools The difference between science and art is one of emphasis Physicians share in man's tendency to dichotomize his problems and ignore the more difficult portion

Experience, kindliness, morality and empathy are consistent with objectivity, knowledge, hard work, intellectual curiosity and a thorough understanding and use of the scientific method in the practice of medicine The seriously ill patient needs both the art and the science of medicine The practicing physician who is provided with great insight and the assurance of a well rounded decision making potentiality will be better able to communicate the art of medicine to his patient The earlier chapters of this book have attempted to bring together a sample of the changes which are occurring in medical practice related to our new ability to encode information and manipulate it with the speed of electricity and in prodigious quantity and complexity

Mathematical techniques have been developing mainly on a theoretical level for centuries At the same time mountains of facts about biology, medicine and the problems of individual patients have continued to accumulate The catalyst that makes possible the solution of medical problems and that also makes possible the application of mathematics to organization of these facts in this solution, is *the stored program digital computer* This is a third and new dimension which not only establishes a triangle in our view but even makes more visible the line which exists between *medicine* and *mathematics*

Here then is the gauntlet The computer is a very complex tool which needs some study This is the easiest part of the problem, however There are those who can help with the use of the tool providing a physician knows his problem and provided he can familarize himself with the pertinent mathematical techniques

There are many relatively simple problems that require little more than grade school arithmetic As a physician works on his chosen problem, however, he will usually soon be led into algebra Depending upon the problem, he will then begin to see the need for symbolic logic or probability and statistics and possibly some special technique in numerical analysis or the calculus

Each step in the learning process increases one's power of creativity The mathematical techniques are very powerful tools but without the computers and application by knowledgeable individuals these tools lie dormant For the practicing physician the required area of study is applied mathematics Most physicians will quickly lose their way if they attempt to become theoretical mathematicians before they attempt to make any use of mathematical skills

Modern algebra such as the recent course presented on television in Continental Classroom is a basic requirement and should help the physician to understand such things as numbers, lines, sets, vectors, matrices and determinants One of the biggest deterrents to the study and understanding of mathematics is notation There is no point in continuing to read beyond one's understanding of the symbols Such notation as the symbols for summation, subscripts, functions, sets, etc must be learned as they initially appear in the text of any book

Multiple texts may contribute to understanding as long as the notation is similar and need not contribute to the confusion if careful watch is made for the differences in notation Programed learning would seem to be a very natural solution to the problem of self-help in developing mathematical skills for the medical profession This is a companion to computer programing and should be watched for, encouraged and made use of to its maximum in medical education

After modern algebra the physician's needs will diverge but not completely nor permanently The initial study of probability and statistics may seem completely different from the study of calculus or symbolic logic, but as the study, and especially the application develops, there is increasing pressure to satisfy needs in other mathematical methods

For the diagnostician symbolic logic or Boolian algebra will seem attractive because it facilitates encoding of information and the sorting and decision making that is required in analyzing diagnostic decision processes

The clinical investigator may find his initial interest greatest in statistics He wants to know the sample average, range and distribution and the tests for significance of results He is interested in experimental design and the types of analytical problems that are illustrated in Chapter VI The basic scientist and clinician find a common area of interest in the use of statistics

Vector algebra, matrices and determinants become increasingly useful as the physician progresses in statistical applications and the more complicated experimental designs, factor analysis and the many varieties of least square analysis

Sooner or later it becomes apparent that the world is very nonlinear, that distributions are skewed and that more versatile approaches are required Relationships and change become of sufficient interest as to demand familiarity with differential equations, calculus and special techniques of curve analysis

The clinician who is in practice may not feel the need for education in mathematics unless he has one or more specific problems which he is interested in trying to solve Given the stimulation of a problem and a little guidance in developing the skills required for solution of that problem, many physicians are picking up the necessary information even though the past mathematical content of premedical and medical education has been appallingly dilute

Clinicians have more obligation to the revolution in automation than simply learning more mathematics Physicians are custodians of the most important biological questions of our time It is necessary for physicians to remain involved in medical research as active participants Physicians must recognize that every individual, regardless of his training, remains ignorant and incompetent in many areas outside of his own It is necessary to establish and maintain the necessary working relationships to solve these problems Physicians must be ready to cross over at least into the fringes of other disciplines and must extend a helpful and a ready hand to those willing to walk on the medical grounds No physician can abdicate his responsibility for the advancement of his own field simply because of the difficulties imposed by the very progress taking place in his own and related fields

The initial period of these new developments will be most exasperating because it requires so much time to develop programs and to instigate the use of these time saving devices However, computers, by performing routine tasks, will make more time available This will lead to other problems which physicians will be required to study

Physicians will need to help plan for the use of time of those less fortunate individuals whose free-time becomes a burden instead of a gift of automation There is argument as to whether automation will make more jobs than it eliminates Certainly it will create problems in adjustment for those unable to adapt to change and it will undoubtedly create some new medical problems There are also social implications of automation which physicians need to prepare for, such as the maintenance of maximum freedom of medical practice and the maintenance of optimum patient physician relationships

Because of the increasing rapidity of change it is now more important that physicians look closely at the educational needs of the generation of medical students entering medical school as well as those successors who are in the lower grades or who will soon be entering school for the first time Are there any real differences in the early educational needs for physicians from the needs of other professional and nonprofessional men? What are the current trends and the educational requirements generally? What will be the world, community, medical and personal problems of tomorrow's physicians? How should tomorrow's physician be educated today?

Acheson, R M Observer Error and Variation in the Interpretation of Electrocardiograms in an Epidemiological Study of Coronary Heart Disease *Brit J Prev Soc Med*, *1960, 14* 99

Adams, S The Medical Literature Crisis—Past and Present *Med Dokum*, 1961, *5* 93

Adey, W R Use of Correlation Analysis in EEG Studies of Conditioning *Electroenceph Clin Neurophysiol*, 1961, *20* 41

Adey, W R Walter, D O, and Hendrix, C E Computer Techniques in Correlation and Spectral Analyses of Cerebral Slow Waves During Discriminative Behavior *Exp Neurol*, 1961, *3* 501

Ainsworth, G C Storage and Retrieval of Biological Information *Nature*, 1961, *191* 12

Alderson, W Computerized Information System for Medical Practice and Research *Circulat Res*, 1962, Part 2, *11* 615

Allen, L E Toward A Procedure for Logically Cataloging Knowledge *Amer Doc*, 1959, *10* 296

Attneave, F *Applications of Information Theory to Psychology* Holt, New York, 1959

Bailey, N T, Jr Statistics in Hospital Planning and Design *J Royal Stat Soc Ser B*, 1956, *5* 146

Bailey, N T, Jr The Mathematical Theory of Epidemics *Heredity*, 1960, *15* 112

Balintfy, J L *A Stochastic Model for the Analysis and Prediction of Admissions and Discharges in Hospitals* Pergamon Press, New York, 1960

Barlow, J S Rhythmic Activity Induced by Photic Stimulation in Relation to Intrinsic Alpha Activity of the Brain in Man *Electroenceph Clin Neurophysiol*, 1960, *12* 317

Barlow, J S Autocorrelation and Crosscorrelation Techniques in EEG Analysis *Electroenceph Clin Neurophysiol*, 1961, *20* 31

Barlow, J S Analysis of Control of Eye Movements in Man by Means of Autocorrelation and Crosscorrelation *15th Annl Conf on Eng in Med and Biology*, 1962, p 4

Barton, J Electronics Is Changing the Structure of Hospitals and Medical Care *Mod Hosp*, 1960, *94* 81

Bartlett, M S The Use of Transformations *Biometrics*, 1947, *3* 39

Baruch, J J Hospital Research and Administration with a Digital Computer *Circulat Res*, 1962, Part 2, *11* 629

Battersby, E J, and Newman, E V Analog Computer Extrapolation of Indicator Dilution Curves *J Clin Invest*, *1962, 41* 1343

Bechtoldt, H P An Empirical Study of the Factor Analysis Stability Hypothesis *Psychometrika,* 1961, *26* 405

Bechtoldt, H P, Benton, A L, and Fogel, M L An Application of Factor Analysis in Neuropsychology *Psychol Rec*, 1962, *12* 147

Bechtoldt, H P Factor Analysis and the Investigation of Hypotheses *Percept and Motor Skills,* 1962, *14* 319

Beckett, P G Computers and Clinical Psychiatry *IRE Trans on Med Electronics,* 1960, ME-7, 248

Beier, W, Brehme, K H, and Glass, K Use of Computers in Biology and Medicine with Special Reference to Biophysical Problems *Z Aerztl Fortbild,* 1962, *56* 271

Bellville, J W, and Seed, J C The Use of Analog Computer for Measurement of Respiration Depression *Trans N Y Acad Sci*, 1959, *22* 34

Bellman, R *Dynamic Programming* Princeton Univ Press, Princeton, 1957

Bellman, R. *Adaptive Control Processes A Guided Tour* Princeton Univ Press, Princeton, 1961

Bellman, R. Mathematical Experimentation and Biological Research *Fed Proc*, 1962, *21* 109

Bennett, W R Methods of Solving Noise Problems *Proc IRE,* 1956, *44* 609

Berman, M, Shahn, E, and Weiss, M. F The Routine Fitting of Kinetic Data to Models, A Mathematical Formalism for Digital Computers *Biophysical Jour*, 1962, *2* 275

Berman, M, Weiss, M F, and Shahn, E Some Formal Approaches to the Analysis of Kinetic Data in Terms of Linear Compartmental Systems *Biophysical Jour*, 1962, *2* 289

Bernard, C *An Introduction to the Study of Experimental Medicine* Dover, New York, 1865

Best, W R Use of a Digital Electronic Computer for Evaluation of Serial Clinical Observations in Acute Leukemia *J Lab Clin Med*, 1961, *58* 799

Best, W R. Use of Digital Electronic Computer for Evaluation of Serial Clinical Observations in Patients with Acute Leukemia *Meth Infor in Med*, 1962, *1* 56

Best, W R The Potential Role of Computers in Medical Practice *J A M A*, 1962, *182* 124

Bickford, R G Scope and Limitations of Frequency Analysis *Electroenceph Clin Neurophysiol*, 1961, *20* 9

Bickford, R G, Poole, E W, and McCarthy, C E Some Applications in EEG Analysis *Fed Proc*, 1962, *21* 103

Bliss, C I *The Statistics of Bioassay with Special Reference to the Vitamins* Academic Press, New York, 1952

Blumberg, M S Hospital Automation, The Needs and the Prospects *Hospitals*, 1961, *35* 34

Blumberg, M S When Should Hospitals Install Computers *Mod Hosp* 1962, *98* 14

Bolie, V W Glucose Insulin Feedback Theory *Proc 3rd Int Conference Med Elect*, 1960, *390* 42

Bonato, R R and Waxman, B D A General Cross Classification Program for Digital Computers *Behav Sci*, 1961, *6* 347

Boole, G *An Investigation of the Laws of Thought* Dover, New York, 1854

Booth, A D Some Applications of Digital Computers in Medicine *Phy Med Biol*, *1962*, *6* 377

Borko, H *Computer Applications in the Behavioral Sciences* Prentice Hall Englewood Cliffs, 1962

Borko, H Computer Simulation of Neurophysiological and Social Systems *Behav Sci*, *1962*, *7* 407

Brandt, E N, Hagans, J A, and Schottstaedt, W W The Computer as a Diagnostic Aid in Medicine *J Okla Med Assoc*, 1962, *55* 213

Brannick, L J An Analog Approach to Computer Diagnosis *IRE Trans on Med Electronics*, 1960, ME-7, 247

Braunstein, J R, Franke, E K, and Ing Autocorrelation of Ventricular Response in Atrial Fibrillation *Circulat Res*, 1961, *9* 300

Brazier, M A B Computer Techniques in EEG Analysis *Electroenceph Clin Neurophysiol*, 1961, Suppl 20

Brazier, M A B Pain Sensory Modality Stimulation Studied by Computer Analysis *Ann N Y Acad Sci*, 1961, *92* 1054

Brazier, M A B The Analysis of Brain Waves *Sci Amer*, *1962* *206* 142

Brody, D A and Romans, W E A Model Which Demonstrates the Quantitative Relations Between the Electromotive Forces of the Heart and the Extremity Leads *Amer Heart J*, 1953, *45* 263

Brody, D A, Bradshaw, J C, and Evans, J W The Elements of an Electrocardiographic Lead Tensor Theory *Bull Math Biophys*, 1961 *23* 31

Brody, D A The Intrinsic Electrocardiographic Properties of Uniform Double Layers *Amer Heart J*, 1962, *63* 841

Brody, D A Use of Computers in Electrophysiology *Circulat Res*, 1962, Part 2, *11* 549

Brodman, K, Erdmann, A J, Lorge, I, and Wolff, H G The Cornell Medical Index Health Questionnaire, As a Diagnostic Instrument *J A M A*, 1951, *145* 152

Brodman, K, Van Woerkom, A J, Erdmann, A J, and Goldstein, L S Interpretation of Symptom with a Data-Processing Machine *A M A Arch Intern Med*, 1959, *103* 776

Brodman, K Diagnostic Decisions by Machine *IRE Trans on Med Electronics*, 1960, ME-7, 216

Brooks, B W Automated Medical Record Systems in Industry *Industr Med Surg*, 1958, *27* 171

Brownell, G L, Cavicchi, R V, and Perry, K E An Electric Analog for Analysis of Compartmental Biological Systems *Rev Sci Instr*, 1953, *24* 704

Bruce, R A and Cox, A R Medical Diagnosis by Computer *Northw Med*, 1961, *60* 1075

Bullock, T H Neuron Doctrine and Electrophysiology *Science*, 1959, *129* 997

Burch, N R Automatic Analysis of the Electroencephalogram, A Review and Classification of Systems *Electroenceph Clin Neurophysiol*, 1959, *11* 827

Bush, V Science in Medicine and Related Fields *Med Ann D C*, 1953, *22* 1

Bushor, W E Medical Electronics, Part II, Diagnostic Systems and Visualization *Electronics*, 1961, *34* 46

Caceres, C A, Kelser, G A, and Mize, W R Formation of the –PR Segment– *Circulation*, 1959, *20* 229

Caceres, C A and Kelser, G A Duration of the Normal P Wave *Amer J Cardiol*, 1959, *3* 449

Caceres, C A Observer Instruments and Computers in Heart Disease Control and as Diagnostic Aids *Proc 2nd IBM Med Symposium*, 1960, Sept 28, 365

Caceres, C A and Rikli, A E The Digital Computer as an Aid in the Diagnosis of Cardiovascular Disease *Trans N Y Acad Sci*, 1961, *23* 240

Caceres, C A, Steinberg, C A, Abraham, S, Carbery, W, McBride, J M, and Tolles, W E Computer Extraction of Electrocardiographic Parameters *Circulation*, 1962, *25* 356

Caceres, C A Integration of Data in Diagnosis *Circulat Res*, 1962, Part 2, *11* 563

Caceres, C A How Can the Waveforms of a Clinical Electrocardiogram Be Measured Automatically by a Computer *IRE Trans on Bio Med Electronics*, 1962, BME-9, 21

Cady, L D Digital Computer Analysis of the Electrocardiogram *IRE Trans on Bio-Med Electronics*, 1961, BME-8, 204

Cady, L D, Gertler, M M, Gottsch, L G, and Woodbury, M A The Factor Structure of Variables Concerned with Coronary Artery Diseases *Behav Sci*, 1961, *6 1*

Cady, L D, Woodbury, M A, Tick, L J, and Gertler, M M A Method for Electrocardiogram Wave-Pattern Estimation Example Left Ventricular Hypertrophy *Circulat Res*, 1961, *9* 1078

Cady, L D Communication with a Computer *Circulat Res*, 1962, Part 2, *11* 497

Camp, G C and Sugarman, H A Operations Research in Hospital Planning *Oper Res*, 1962, *10* 55

Campbell, J D and Caron, H S Data Processing by Optical Coincidence *Science*, 1961, *133* 1333

Carbery, W J, Steinberg, C A, Tolles, W E and Freiman, A H Automatic Methods for the Analysis of Physiological Data *Aerospace Med*, 1961, *32* 52

Caster, W O *Use of a Digital Computer in the Study of Eating Habit* Patterns *Amer J Clin Nutr*, 1962, *10* 98

Chance, B and Williams, G R Techniques of Application of Computers *Advance Enzymol*, 1956, *17* 65

Chance, B Analogue and Digital Representations of Enzyme Kinetics *J Biol Chem*, 1960, *235* 2440

Chance, B, Higgins, J J, and Garfinkel, D Analogue and Digital Computer Representations of Biochemical Processes *Fed Proc*, 1962, *21* 75

Charnes, A, Cooper, W W, and Henderson, A *Introduction to Linear Programming* J Wiley and Sons, New York, 1953

Cheshier, R G Machine Information Search System *Bull Med Libr Assoc*, 1962, *50* 481

Christ, C F, Hildreth, C, Liu, T C, and Klein, L T *Symposium on* Simultaneous Equation Estimation *Econometrica*, 1960, *28* 835

Churchman, C W, Ackoff, R L, and Arnoff, E L *Introduction to Operations Research* J Wiley and Sons, New York, 1957

Claringbold, P J Orthogonal Contrasts in Slope Ratio Investigations *Biometrics*, 1959, *15* 307

Clauss, R H, Missier, P, Reed, G E, and Tice, D Assisted Circulation by Counter-Pulsation with an Intra Aortic Balloon Methods and Effects *15th Annl Conf on Eng in Med and Biology*, 1962, p 44

Clynes, M Respiratory Sinus Arrhythmia Laws Derived from Computer Simulation *J Appl Physiol*, *1960, 15 863*

Clynes, M Computer Analysis of Reflex Control and Organization Respiratory Sinus Arrhythmia *Science*, 1960, *131* 300

Clynes, M Use of Computers for Physiological Discovery and for Diagnosis by Dynamic Simulation *Circulat Res*, 1962, Part 2, *11* 515

Clynes, M The Non Linear Biological Dynamics of Unidirectional Rate Sensitivity Illustrated by Analog Computer Analysis, Pupillary Reflex to Light and Sound, and Heart Rate Behavior *Ann N Y Acad Sci*, 1962, *98* 806

Cochran, W G The Role of Mathematics in the Medical Sciences *New Engl J Med*, 1961, *265* 176

Cochran, W G and Hopkins, C E Some Classification Problems with Multivariate Qualitative Data *Biometrics*, 1961, *17* 10

Connor, R J, Flagle, C D, Hsieh, R K C, Preston, R A, and Singer, S Effective Use of Nursing Resources A Research Report *Hospitals*, 1961, *35* 30

Cope, F W An Elastic Reservoir Theory of the Human Systemic Arterial System Using Current Data on Aortic Elasticity *Bull Math Biophys*, 1960, *22* 19

Cope, F W A Method for the Computation of Aortic Distensibility in the Living Human Patient and Its Use for the Determination of the Aortic Effects of Aging, Drugs, and Exercise *Bull Math Biophys*, 1961, *23* 337

Cope, F W A Method for Calculation of Human Systolic and Diastolic Blood Pressure Using an Elastic Reservoir Theory of the Systemic Arterial System, and Some Clinical and Physiological Applications *Bull Math Biophys*, 1962, *24* 137

Cornfield, J Joint Dependence of Risk of Coronary Heart Disease of Serum Cholesterol and Systolic Blood Pressure A Discriminant Function Analysis *Fed Proc*, 1962, *21* 58

Cosma, J, Stallmann, F W, and Pipberger, H V Digital Computation of the Spatial Ventricular Gradient in Normal and Premature Ventricular Contraction and Intermittent Bundle Branch Block *Circulation*, 1961, *24* 910

Cox, D R and Smith, W L *Queues* J Wiley and Sons, New York, 1961

Curtiss, J H On Transformations Used in the Analysis of Variance *Ann Math Stat*, 1943, *14* 107

Dantzig, G B and DeHaven, J C A Math Model of the Human External Respiratory System *Biol Med*, 1961, *4* 324

Dantzig, G B A Mathematical Model of the Human External Respiratory System *Perspect Biol Med*, 1961, *4* 324

Davenport, W B and Root, W L *An Introduction to the Theory of Random Signals and Noise* McGraw-Hill, New York, 1958

Davies, L G Observer Variation in Reports on Electrocardiograms *Brit Heart J*, *1958*, *20* 153

Davis, G Computer Clicks with Accounting Office *Mod Hosp*, 1962, *99* 113

Davis, P A Technique and Evaluation of the Electroencephalogram *J Neurophysiol*, 1941, *4* 92

Dawber, T R, Kannel, W B, Love, D E, and Streeper, R. B The Electrocardiogram in Heart Disease Detection *Circulation*, 1952, *5* 559

Dawber, T R and Kannel, W B An Epidemiologic Study of Heart Disease The Framingham Study *Nutr Rev*, 1958, *16* 1

Dawber, T R, Kannel, W B, Revotskie, N, Stokes, J, Kagan, A, and Gordon, T Some Factors Associated with the Development of Coronary Heart Disease *Amer J Public Health*, 1959, *49* 1349

Dawber, T R and Kannel, W B Computers in Epidemiologic Research Uses in the Framingham Study *Circulat Res*, 1962, Part 2, *11* 587

Dawson, G D A Summation Technique for the Detection of Small Evoked Potentials *Electroenceph Clin Neurophysiol*, 1954, *6* 65

Dayhoff, M O and Ledley, R S Use of Computers in Protein Biochemistry *15th Annl Conf on Eng in Med and Biology*, 1962, p 34

Dedijer, S Measuring the Growth of Science *Science*, 1962, *138* 781

Defares, J G and Sneddon, I N *An Introduction to the Mathematics of Medicine and Biology* North Holland, Amsterdam, 1960

Defares, J G, Derksen, H E, and Duyff, J W Cerebral Blood Flow in the Regulation of Respiration *Acta Physiol Pharmacol Neerl*, *1960*, *9* 327

Deland, E C Simulation of a Biological System on an Analog Computer *IRE Trans Elec Comput*, 1962, EC-11, 17

Demorr, P and Meulepas, E Disposal of Plasma Corticoids During Adrenocorticotropic Hormone Infusion as Studied by Analog Computer Analysis *J Clin Endocr*, *1962*, *22* 544

Devaux, P An Electronic Computer in the Service of Scholars *Geneesk Gids*, 1960, *38* 97

Diamond and Solomon *Information and Error* Basic Books, Inc, New York, 1959

Diamond, E L and Lilienfeld, A M Effects of Errors in Classification and Diagnosis in Various Types of Epidemiological Studies *Amer J Public Health*, 1962, *52* 1137

Dietrich, E V Machine Retrieval of Pharmacological Data *Science*, 1960, *132* 1556

Dornhorst, A C and Lee, G D Analog Computer for Calculating the Alveolar-Capillary Oxygen Gradient *J Appl Physiol*, 1957, *11* 331

Doyle, J T, Dawber, T R, Kannel, W B, Heslin, A S and Kahn, H A Cigarette Smoking and Coronary Heart Disease *New Engl J Med*, 1962, *266* 796

Dryer, B V Thinking Men and Thinking Machines in Medicine *J Med Educ*, 1963, *38* 82

Duncan, A J *Quality Control and Industrial Statistics* R D Irwin, Inc, Homewood, Ill, 1959

Easterday, O, Mumford, P and Potter, L Use of Computers for Statistically Analyzing Toxicity Data *Fed Proc*, 1960, *19* 391

Eblad, R and Lane, R Digital Computers and Medical Logic *IRE Trans on Med Electronics*, 1960, ME-7, 283

Eccles, J C The Physiology of Imagination *Sci Amer*, 1958, *199* 135

Eden, M Storage and Retrieval of the Results of Clinical Research *IRE Trans on Med Electronics*, 1960, ME-7, 265

Einthoven, W Le Telecardiogramme *Arch Int Physiol*, 1906, *4* 132

Einthoven, W, Fahr, G and DeWaart, A On the Direction and Manifest Size of the Variations of Potential in the Human Heart and on the Influence of the Position of the Heart on the Form of the Electrocardiogram *Amer Heart J*, 1950, *40* 163

Eisen, J N A Note on the Application of Orthogonal Polynomials to Data on Palatability When the Levels of a Treatment are Unequally Spaced *Food Tech*, 1960, *14* 663

Elston, R C On Additivity in the Analysis of Variance *Biometrics*, 1961, *17* 209

Engel, L L Application of Automation to Urinary Steroid Assays *Science*, 1962, *136* 160

Enslein, K Applications of a Method of Pattern Recognition to Biomedical Problems *15th Annl Conf on Eng in Med and Biology*, 1962, p 35

Epstein, F H, Doyle, J T, Pollack, A A, Pollack, H, Robb, G P and Simonson, E Observer Interpretation of Electrocardiograms Suggestions for Objective Interpretations *J A M A*, 1961, *175* 847

Ezekiel, M and Fox, K A *Methods of Correlation and Regression Analysis* J Wiley and Sons, New York, 1959

Fairweather, D V I and Millar, D G Electronic Digital Computer A Tool for Research *Lancet*, 1962, *2* 1214

Farley, B G and Clark, W A Simulation of Self Organizing Systems by Digital Computer *Trans IRE*, 1954, PGIT-4, 76

Farley, B G Recognition of Patterns in the EEG *Electroenceph Clin Neurophysiol*, 1961, *20* 49

Farley, B G Some Results of Computer Simulation of Neuron Like Nets *Fed Proc*, 1962, *21* 92

Farley, B G, Frishkopf, L S, Clark, W A and Gilmore, J T Computer Techniques for the Study of Patterns in the Electroencephalogram *IRE Trans on Bio Med Electronics*, 1962, BME 9, 4

Farmer, J E Automated Programming for Multinomial Distribution Problems on the IBM 650 *Behav Sci*, 1961, *6* 164

Farrer, J T Use of a Digital Computer in the Analysis of Intestinal Motility Records *IRE Trans on Med Electronics*, 1960, ME-7, 259

Farrow, R L and Stacy, R W Harmonic Analysis of Aortic Pressure Pulses in the Dog *Circulat Res*, 1961, *9* 395

Faulconer, A and Bickford, R G *Electroencephalography in Anesthesiology* Thomas, Springfield, 1960

Feder, D P Automatic Lens Design with a High Speed Computer *J Optic Soc Amer*, 1962, *52* 177

Feigenbaum, E A A Simulation of Verbal Learning Behavior *Proc Western Joint Comptr Conf* ACM AIEE-IRE, 1961, 121

Feigenbaum, E A An Experimental Course in Simulation of Cognitive Processes *Behav Sci*, 1962, *7* 244

Fein, L The Computer Related Sciences (Synnoetics) at a University in the Year 1975 *Amer Scient*, 1961, *49* 149

Feldman, H Programmed Instruction What It Is How Hospitals Can Use it *Mod Hosp*, 1962, *99* 75

Feller, W *An Introduction to Probability Theory and Its Application* J Wiley and Sons, New York, 1960

Felter, R F and Thompson, J D Computer Models for the Planning and Utilization of Hospital Facilities *Oper Res*, 1962, *10* 63

Fish, B R Applications of an Analog Computer to Analysis of Distribution and Excretion Data *Health Physics*, 1958, *1* 276

Fisher, R A *Statistical Methods and Scientific Inference* Hafner, New York, 1959

Fisher, R A *The Design of Experiments* Hafner, New York, 1960

Fisher, R A and Yates, F *Statistical Tables for Biological, Agricultural and Medical Research* Hafner, New York, 1962

Flagle, C D *Operations Research in a Hospital* Johns Hopkins Press, Baltimore, 1960

Flagle, C D *The Problem of Organization for Hospital Inpatient Care* Pergamon Press, New York, 1960

Flagle, C D *Operations Research in System Engineering* Chapter 25, Operations Research in a Hospital Johns Hopkins Press, Baltimore, 1960

Flagle, C D Operations Research in the Health Services *Circulat Res*, 1962, Part 2, *11* 621

Flores, I *Computer Logic* Prentice Hall, Englewood Cliffs, 1960

Forsyth, G and Logan, R F L *The Demand for Medical Care* Oxford Univ Press, London, 1961

Franke, E K, Braunstein, J R and Zellner, D C Study of High Frequency Components in Electrocardiogram by Power Spectrum Analysis *Circulat Res*, 1962, *10* 870

Fraser, A S Simulation of Genetic Systems by Automatic Digital Computers Effects of Reproductive Rate, and Intensity of Selection, on Genetic Structure *Aust J Biol Sci*, 1960, *13* 344

Freeman, H *Finite Differences for Actuarial Students* Cambridge Univ Press, Cambridge, 1960

Freeman, W J Simulation of Paleocortical Evoked Potential with Linear Second Order Equation *Fed Proc*, 1962, *21* 356

French, C S and Fork, D C Computer Solutions for Photosynthesis Rates From a Two Pigment Model *Biophysical Jour*, 1961, *1* 669

Fry, D L The Use of Computers in Physiologic Diagnosis *IRE Trans on Med Electronics*, 1960, ME 7, 269

Fuknda, B and Sugita, M Mathematical Analysis of Metabolism Using an Analog Computer Isotope Kinetics of Iodine Metabolism in the Thyroid Gland *J Theor Biology*, 1961, *1* 440

Gaddum, J H Lognormal Distributions *Nature*, 1945, *156* 463

Galambos, J T and Cornell, R G Mathematical Models for the Study of the Metabolic Pattern of Sulfate *J Lab Clin Med*, 1962, *60* 53

Gardner, R M Analog Computer Simulates Heart Response to Nerve Simulation *Elec Engr*, 1961, *80* 979

Garfinkel, D Computer Simulation of Steady-State Glutamate Metabolism in Rat Brain *J Theor Biology*, 1962, *3* 412

Garland, L H Studies on the Accuracy of Diagnostic Procedures *Am J Roentgenol*, 1959, *82* 25

Garn, S M Automation in Anthropometry *J Phys Anthropol*, 1962, *20* 387

Garrett, E R Psicofuranine Kinetics and Mechanisms In Vivo with *the Application of the Analog Computer* *J Pharmacol Exp Therap*, 1960, *130* 106

Garry, H *Applications in Continuous Averaging of Medicine* *J Med Electr*, 1961, Nov 21

Geary, R C *Moments of the Ratio of the Mean Deviation to the* Standard Deviation for Normal Samples *Biometrika*, 1936, *28* 295

Geisler, C D, Frishkopf, L S and Rosenblith, W A Extracranial Responses to Acoustic Clicks in Man *Science*, 1958, *128* 1210

Geisler, C D, Molnar, C E, Peake, W T, Weiss, T F and Sandel, T T Loudness, Judgments on Clicks *J Acoust Soc Am*, 1959, *31* 128

Gerbarg, D S, Holcomb, F W, Hofler, J J, Bading, C E, Schultz, G L and Sears, R. E *Analysis of Phonocardiogram by a Digital* Computer *Circulat Res*, 1962, Part 2, *11* 569

Gertler, M M, Driskell, M M, Bland, E F, Garn, S M, Lerman, J and Levine, S A Clinical Aspects of Coronary Heart Disease An *Analysis of 100 Cases in Patients 23 to 40 Years of Age with Myocardial* Infarction *J A M A*, 1951, *146* 1291

Gertler, M M, Garn, S M and White, P D Young Candidates for Coronary Heart Disease *J A M A*, 1951, *147* 621

Geselowitz, D B *Multipole Representation for an Equivalent Cardiac* Generator *Proc IRE*, 1960, *48* 75

Geselowitz, D B The Anatomy and Function of the Computer *Circulat Res*, 1962, Part 2, *11* 489

Gibbs, F A. and Grass, A M Frequency Analysis of Electroencephalograms *Science*, 1947, *105* 132

Gimenez, J L, Carter, B L, Stewart, G H and Lynch, P R Flow Pattern Analysis of Particulate Contrast Material in Blood Vessels Experimental Cinefluorography at 260 Frames/sec, Superimposed oscillography of Cardiovascular Events and Automatic Data Reduction *Acta Radiol*, 1961, *55* 350

Goff, K W An Analog Electronic Correlator for Acoustic Measurements *J Acoust Soc Am*, 1955, 27 223

Goldstein, H Problems of the Statistician in Collaborative Research. *Public Health Rep*, 1961, *76* 43

Goldstein, M H and Kiang, N Y Synchrony of Neural Activity in Electric Responses Evoked by Transient Acoustic Stimuli *J Acoust Soc Am*, 1958, *30* 107

Goldstein, M H, Kiang, N Y and Brown, R M Responses of the Auditory Cortex to Acoustic Stimuli *J Acoust Soc Am*, 1959, 31 356

Gorby, J H Electronic-Brain—Handles Patient Records *Hosp Topics*, 1959, *37* 45

Gordon, T, Moore, F E, Shurtleff, D and Dawber, T R Some Methodologic Problems in the Long-Term Study of Cardiovascular Disease Observations on the Framingham Study *J Chron Dis*, 1959, *10* 186

Gouras, P, Armington, J C, Kropfl, W, Tepas, D I and Gunkel, R D Evaluation of Visual Function in Man by Electronic Computation of Evoked Potentials *Rochester Conf Data Acquis and Process Biol and Med*, 1962

Graybiel, A, McFarland, R A, Gates, D C and Webster, F A Analysis of the Electrocardiograms Obtained from 1,000 Young Healthy Aviators *Amer Heart J*, 1944, *27* 524

Gripe, A H Pitfalls in the Electrocardiographic Diagnosis of Left Ventricular Hypertrophy A Correlative Study of 200 Autopsied Patients *Circulation*, 1959, *20* 30

Guest, P G *Numerical Methods of Curve Fitting* Cambridge Univ Press, Cambridge, 1961

Guillemin, E A *The Mathematics of Circuit Analysis* J Wiley and Sons, New York, 1949

Gulliksen, H Mathematical Solutions for Psychological Problems *Amer Scient*, 1959, *47* 178

Gustafson, J E, Daniel, C M, White, W J and Staples, L F Computer Diagnosis of Ventricular Hypertrophy in Children *Rochester Conf Data Acquis and Process Biol and Med*, 1962

Hagans, J A and Schottstaedt, W N Computers in Modern Medicine *J Okla Med Assoc*, 1962, *55* 96

Hagan, W and Horowitz, M Estimate of Information in an Electrocardiac Signal *15th Annl Conf on Eng in Med and Biology*, 1962, p 47

Hamming, R W Controlling the Digital Computer *Sci Mo*, 1957, Oct, 169

Hara, H H Analog Simulation of Human Blood Circulation *Proc San Diego Symp Biomed Eng*, 1962, 107

Harken, D E and Lefemine, A A Myocardial and Circulatory Assistance by Synchronized Counterpulsation *15th Annl Conf on Eng in Med and Biology*, 1962, p 42

Harmon, H H *Modern Factor Analysis* Univ of Chicago Press, 1960

Harrington, R W Patient Data A Computer Based System *Hosp Progr*, 1962, *43* 64

Hayne, R L and Balthaser, J F The Evaluation of Clinical Data with the IBM 101 *Med Dokum*, 1960, *4* 61

Heaps, H S An Iterative Method for Obtaining Statistical Frequency Distributions *Nature*, *1962*, *196* 1236

Heinmets, F and Herschman, A Quantitative Analysis of Metabolic Processes *Phys Med Biol*, *1960*, *4* 238

Hemphill, E C The Probable Impact of the NIH Training Grant Program on the Future Supply of Biostatisticians *Amer J Public Health*, 1961, *51* 1775

Henkin, L Are Logic and Mathematics Identical *Science*, 1962, *138* 788

Hidalgo, J U, Nadler, S B and Bloch, T The Use of the Electronic Digital Computer to Determine Best Fit of Blood Volume Formulas *J Nucl Med*, 1962, *3* 94

Hoen, K and Grandage, A H E Calculation of Inbreeding in Family Selection Studies on the IBM 650 Data Processing Machine *Biometrics*, 1960, *16* 292

Hoffman, I A New Approach to Processing Electrocardiograms *J A M A*, 1962, *181* 130

Hogg, R V and Craig, A T *Introduction to Mathematical Statistics* MacMillan Co, New York, 1959

Hollingsworth, T H Using an Electronic Computer in a Problem of Medical Diagnosis *J Royal Stat Soc Ser* A 1959, *122* 221

Hommes, F A The Integrated Michaelis Menten Equation *Arch Biochem Biophys*, 1962, *96* 28

Hommes, F A Analog Computer Studies of the Carboxypeptidase A Catalyzed Hydrolysis of Chloroacetyl—L Phenylalanine *Arch Biochem Biophys*, 1962, *96* 37

Hommes, F A Analog Computer Studies of a Simple Enzyme Catalyzed Reaction *Arch Biochem Biophys*, 1962, *96* 32

Hopkins, R E and Spencer, G Creative Thinking and Computing Machines in Optical Design *J Optic Soc Amer*, 1962, *52* 172

Horgan, J D and Lange, R L Simulation of the Chemical and Circulatory Factors in the Dynamic Analysis of the Human Respiratory Control System *15th Annl Conf on Eng in Med and Biology*, 1962, p 53

Horst, P *Matrix Algebra for Social Scientists* Holt, Rinehart & Winston New York, 1963

Horvath, W J Applicability of the Lognormal Distribution to Servicing Times in Congestion Problems. *Oper Res*, 1959, *7* 127

Hough, P V C General Purpose Visual Input for a Computer *Ann N Y Acad Sci*, 1962, *2* 323

Hugenholtz, P G, Forkner, C E and Levine, H D Clinical Appraisal of the Vectorcardiogram in 207 Cases of Myocardial Infarction (Frank System) *Circulation*, 1960, *22* 765

Janaske, P C Information Handling and Science Information a Selected Bibliography 1957-1961 *Amer Inst of Biological Sciences*, Washington, D C, 1962

Jeenel, J *Programming for Digital Computers* McGraw-Hill, New York, 1959

Johnson, D G A Multifactor Method of Evaluating Medical School Applicants *J Med Educ*, 1962, *37* 656

Johnston, F D, McFee, R and Bryant, J An Integration Circuit for Measurement of the Area of the Waves in the Electrocardiogram *Circulation*, 1950, *2* 5

Jones, G M and Milsum, J H On the Physiological Control System for Ocular Stabilization *15th Annl Conf on Eng in Med and Biology*, 1962, p 51

Jones, R W, Li, C C, Meyer, A U and Pinter, R. B Pulse Modulation in Physiological Systems, Phenomenological Aspects *IRE Trans on BIO MED Electronics*, 1961, BME-8, 59

Jones, R. W, Green, D G and Pinter, R. B Mathematical Simulation of Certain Receptor and Effector Organs *Fed Proc*, 1962, *21* 97

Jones, W B and Griffin, J B Evaluation of a Computer Method of Obtaining Aortic Blood Velocity Curves from the Aortic Pulse *Clin Res*, 1961, *9* 16

Jonnard, R A Random Selection System for Automatic Dynamic Biochemical Analysis *Ann N Y Acad SC L*, 1960, *87* 908

Jordan, L R and Hoffman, R C Machines Make Light Work of Medical Data *Mod Hosp*, 1960, *95* 67

Kagan, A, Gordon, T, Kannel, W B and Dawber, T R Blood Pressure and Its Relation to Coronary Heart Disease in the Framingham Study *Hypertension*, 1959, *7* 53

Kagan, A, Dawber, T R, Kannel, W B and Revostkie, N The Framingham Study A Prospective Study of Conorary Heart Disease *Fed Proc*, 1962, *21* 52

Kahn, H A Use of Computers in Analyzing Framingham Data *Circulat Res*, 1962, Part 2, *11* 585

Kane, S H Administrative Significance of Computerized Medical Studies *Circulat Res*, 1962, Part 2, *11* 647

Kiang, N Y The Use of Computers in Studies of Auditory Neurophysiology *Trans Amer Acad Ophthal Oto*, 1961, *65* 735

Kinsey, V E Computer and Eye Research *Med. Electr News*, 1962, *2* 3

Kirkwood, J G and Fitts, D D Statistical Mechanics of Transport

Processes XIV Linear Relations in Multicomponent Systems *J Chem Phys*, 1960, *33* 1317

Klett, S and Sewall, L Automatic Data Processing of Patient's Records *Mental Hosp*, 1961, *12* 26

Korn, G A and Korn, T M *Mathematical Handbook for Scientists and Engineers* McGraw-Hill, New York, 1961

Kornfeld, H Machine Diagnosis European Experiences *Deutsch Med J*, 1961, *12* 642

Krasnoff, S O, Echikson, E and Oscherwitz, M Acute Myocardial Infarction Feasibility of Utilizing Electronic Data-Processing Apparatus to Investigate a Large Series of Cases *Circulat Res*, 1962, Part 2, *11* 591

Krueger-Thiemer, E, Diller, W and Buenger, P Dosage Computation with Program-Directed Computers The Experimental Evaluation of New Sulfanilamides as a Preparation for their Clinical Testing *Klin Wschr*, 1962, *40* 698

Kunz, K S *Numerical Analysis* McGraw-Hill, New York, 1957

Langner, P H and Geselowitz, D B Characteristics of the Frequency Spectrum in the Normal Electrocardiogram and in Subjects Following Myocardial Infarction *Circulat Res*, 1960, *8* 577

Langner, P H and Geselowitz, D B First Derivative of the Electrocardiogram *Circulat Res*, 1962, *10* 220

Lawrence, J N *Puqfua, An IBM 704 Code for Computing Plutonium* Body Burdens *Health Physics*, 1962, *8* 61

Ledley, R S and Lusted, L B Reasoning Foundations of Medical Diagnosis Symbolic Logic, Probability, and Value Theory Aid Our Understanding of How Physicians Reason *Science*, 1959, *130* 9

Ledley, R. S and Lusted, L B Use of Electronic Computers To Aid In Medical Diagnosis *Proc Ire*, 1959, *47* 1970

Ledley, R S Digital Electronic Computers In Biomedical Science *Science*, 1939, *130* 1225

Ledley, R S and Lusted, L B The Role Of Computers In Medical Diagnosis *Med Dokum*, 1961, *5* 70

Ledley, R. S and Lusted, L B Biomedical Electronics Potentialities and Problems *Science*, *1962*, *135 198*

Ledley, R. S National Health Computing System. *Oper Res*, 1962, *10 64*

Lefemine, A A, Lunzer, S and Harken, D Assisted Circulation by Veno-Arterial By-Pass *J A M A*, 1962, *182* 148

Leitch, I Nutrition A Scheme for Classification and Coding of Information for Machine Retrieval *Nutr Abstr Rev*, 1961, 31

Leitch, I Machine Retrieval of Knowledge of Nutrition *Nutr Rev*, 1962, *20* 65

Levick, W R, Bishop, P O, Williams, W O and Lampard, D G Probability Distribution Analyser Programmed for Neurophysiological Research *Nature*, 1961, *192* 629

Lew, E A Biostatistical Pitfalls in Studies of Atherosclerotic Heart Disease *Fed Proc*, 1962, *21* 62

Lewis, B M, Adhikari, P K, Boushy, Sr and Sakamoto, A The Diffusing Capacity of the Lungs in Emphysema Analog Computer Analysis *J Lab Clin Med*, 1961, *58* 938

Lewis, B M, Adhikari, P K, Boushy, S R and Sakamoto, A Analog Computer Analysis of Diffusing Capacity Measurements *J Clin Invest*, 1962, *41* 1379

Lewin, K *Field Theory in Social Science* Harper & Bros, New York, 1951

Lidwell, O M Sepsis in Surgical Wounds Multiple Regression Analyses Applied to Records of Post-Operative Hospital Sepsis *J Hyg*, 1961, *59* 259

Liebrich, K. G and Schwarzfischer, F On Population Genetics of ABO Blood Groups Description of Programs for the Calculation of Gene Frequencies with the Aid of the Electric Magnetic Drum Computer IBM 650 *Blut*, 1960, *6* 134

Lindquist, E F *Design and Analysis of Experiments in Psychology and Education* Houghton Mifflin Company, Boston, 1953

Lindgren, F T, Tandy, R K, Wills, R D and Nichols, A V Computer Analysis of the Blood Fatty Acids *Circulat Res* 1962, Part 2, *11* 577

Lingoes, J C Information Processing in Psychological Research *Behav Sci*, 1962, 7 412

Lipkin, M and Hardy, J D Differential Diagnosis of Hematologic Diseases Aided by Mechanical Correlation of Data *Science*, 1957, *125* 551

Lipkin, M and Hardy, J D Mechanical Correlation of Data in Differential Diagnosis of Hematological Diseases *J A M A*, 1958, *166* 113

Lipkin, M, Engle, R. L, Davis, B J, Zworykin, V K., Ebald, R and Sendrow, M Digital Computer as Aid to Differential Diagnosis Use in Hematological Diseases *A M A Arch Intern Med*, 1961, *108* 56

Lipkin, M Digital and Analogue Computer Methods Combined to aid in the Differential Diagnosis of Hematological Diseases *Circulat Res*, *1962, Part 2, 11* 607

Lotka, A J *Elements of Mathematical Biology* Dover, New York, 1956

Lowell, E L, Williams, C T, Ballinger, R. M and Alvig, D P Measurement of Auditory Threshold with a Special Purpose Analog Computer *J Speech Hearing Res*, 1961, *4* 105

Lowe, T E Application of Computers to Biology and Medicine *Aust J SCI* 1960 *21* 441

Lu, K H and Savara, B S A Statistical Analysis of Relative importance in Factors Affecting Weight Changes in Children From Three to Eight Years Old *Hum Biol*, 1962, *34* 1

Luebs, H Automation at the Front Desk *Hosp Manage*, 1960, *90* 54

Luhn, H P *Keyword In Context Index for Technical Literature* (KWIC Index) IBM Adv Syst Dev Div, Yorktown Heights, N Y, 1959

Lurie, W The Impertinent Questioner The Scientists Guide to the Statistician's Mind *Amer Scient*, 1958, *46* 57

Lusted, L B and Ledley, R S *Mathematical Models in Medical Diagnosis* *J Med Educ*, 1960, *35* 214

Lusted, L B Logical Analysis in Roentgen Diagnosis *Radiology*, 1960, *74* 178

Lusted, L B Computer Programming of Diagnostic Tests *IRE Trans on Med Electronics*, 1960, ME-7, 293

Lusted, L B Application of Computers in Diagnosis *Circulat Res* 1962, Part 2, *11* 599

MacDonald, J Q, Perry, E G, Madison, L L and Seldin, D W An Electric Analogue for Analysis of Tracer Distribution Kinetics in Biological Systems *Radiat Res*, 1957, *6* 585

Magoun, H W *The Waking Brain* Thomas, Springfield, 1958

Maranga, A J Automation Brings Boom to Special Surgery *Mod Hosp*, *1961*, *96* *88*

Martinek, J, Yeh, G C K and Carmine, R. A New System for Electrocardiographic Recording, Analysis and Diagnosis *IRE Trans on Med Electronics*, 1959, ME-6, 112

Mason, E E, Chernigoy, F and Cusminsky, B Computer Analysis in Development of New Diagnostic Methods *J A M A*, 1961, *178* 110

Mason, E E and Chernigoy, F Automatic Data Processing and Experimental Design in Medical Research. *Rochester Conf Data Acquis, and Process Biol and Med*, 1962

Matin, L Fourier Treatment of Some Experiments in Visual Flicker *Science*, 1962, *136* 983

McCormack, J E Changing Concepts in Patient Care *Hosp Progr*, 1962, *43* 57

McFee, R A Trigonometric Computer with Electrocardiographic Applications *Rev Sci Instr*, 1950, *21* 420

McFee, R. and Parungao, A An Orthogonal Lead System for Clinical Electrocardiography *Amer Heart J*, 1961, *62* 93

McGinn, L C The Purpose of a Computer *Circulat Res*, 1962, Part 2, *11* 495

Meneely, G R Some Reflections on Medical Diagnosis by Electronic Data Processing Machines *IRE Trans on Med Electronics*, 1960, Me-7, 309

Meneely, G R, Ball, C O T, Ferguson, J L, Payne, D D, Lorimer, A R and Weiland, R L Use of Computers in Measuring Body Electrolytes by Gamma Spectrometry *Circulat Res* 1962, Part 2, *11* 539

Metcalf, D R A Flexible IBM System for Collating EEG and Clinical Data *Electroenceph Clin Neurophysiol*, 1960, *12* 202

Morris, J P and McKinney, S Mechanical Disease Indexing—A Fast and Efficient System *J A M A*, 1959, *171* 93

Morrison, R T Some Applications of Statistics to Medical Research *Canada Med Assoc J*, 1961, *84* 435

Mosteller, F, Rourke, R E K and Thomas, G B *Probability and Statistics* Addison Wesley, Reading, Mass, 1961

Mueller, C G Numerical Transformations in the Analysis of Experimental Data *Psychol Bull*, 1949, *46* 198

Murtaugh, J S and Payne, G L Communication in the Bio medical Sciences *J Med Educ*, 1962, *37* 1169

Myers, R S Why Automation Can't Replace the Doctor *Mod Hosp*, 1961, *97* 120

Nanus, B The Use of Electronic Computers for Information Retrieval *Bull Med Libr Assoc*, 1960, *48* 278

Nash, F A Differential Diagnosis An Apparatus to Assist the Logical Facilities *Lancet*, 1954, *1* 874

Nash, F A The Mechanical Conservation of Experience, Especially in Medicine *IRE Trans on Med Electronics*, 1960, ME-7, 240

Neel, R B and Olson, J S Use of Analog Computers for Simulating the Movements of Isotopes in Ecological Systems *U S Atom Energy Comm*, 1962

Neisser, U The Imitation of Man by Machine *Science*, 1963, *139* 193

Newell, A and Simon, H A Computer Simulation of Human Thinking A Theory of Problem Solving Expressed as a Computer Program Permits Simulation of Thinking Processes *Science*, 1961, *134* 2011

Newman, E V, McGovern, J F and Arnold, T G The Important Recent Advances in Electrocardiography *A M A Arch Intern Med*, 1955, *96* 591

Noble, F W, Hayes, J E and Eden, M Repetition Analog Computer for Analysis of Sums of Distribution *Proc IRE* 1959, *47* 1952

Noe, F E Computer Analysis of Curves from an Infrared Carbon Dioxide Analyzer and Screen Type Airflow Meter *J Appl Physiol*, 1963, *18* 149

Ornstein, G N The Automatic Analog Determination of Human Transfer Function Coefficients *4th Intl Conf Med Electronics*, 1961, July 16

Orvis, H H, Thomas, R E, Fawal, I A and Evans, J M Serum Cholesterol in Normal Males Versus Males with Coronary Heart Disease Differences With Respect to Age *Amer J Med Sci*, 1961, *241* 167

Osborn, J, Russi, M, Sale, A, Bramson, M L and Gerbode, F Diastolic Augmentation by External Pulsed Pressure *15th Annl Conf on Eng in Med and Biology*, 1962, p 44

Osburn, J O Analog Computer Instruction A Plugboard Teaching Aid *J Chem Educ*, 1961, *38* 492

Overall, J E and Williams, C M *Models for Medical Diagnosis Factor* Analysis *Med Dokum*, 1961, *5 51*

Overall, J E and Williams, C M Models for Medical Diagnosis *Behav Sci*, 1961, *6* 134

Overall, J E and Williams, C M A Computer Procedure for the Diagnosis of Thyroid Function *Rochester Conf Data Acquis and Process Biol and Med*, 1962

Overall, J. E. and Williams, C M. Conditional Probability Program for Diagnosis of Thyroid Function *J A M.A*, 1963, *183 307*

Pace, W H An Analog, Computer Model for the Study of Water and Electrolyte Flows in the Extracellular and Intracellular Fluids *IRE Trans On Bio-Med Electronics*, 1961, BME-8, 29

Pap, A Mathematics, Abstract Entities, and Modern Semantics *Sci Mo*, 1957, *July* 29

Parrish, D, Hayden, T, Garrett, W and Huff, R L Analog Computer Analysis of the Flow Characteristics and Volume of the Pulmonary Vascular Bed *Circulat Res*, 1959, *7* 746

Paycha, F Diagnosis, Therapeutics, Prognosis, and Computers *IRE Trans on Med Electronics*, 1960, ME-7, 288

Peake, W T, Goldstein, M H and Kiang, N Y Steady-State-Auditory Nerve Potentials for Different Stimulus Repetition Rates *J Acoust Soc Am*, 1959, *31* 124

Pearson, E S A Further Development of Tests for Normality *Biometrika*, 1930, *22* 239

Peterson, L H, Jensen, R E and Parnell, J Mechanical Properties of Arteries in Vivo *Circulat Res* 1960, *8* 622

Peterson, L H Introduction to the Principles of Digital and Analogue Computers *Fed Proc*, 1962, *21* 69

Phillips, W, Gorwitz, K and Bahn, A K Electronic Maintenance of Case Registers *Public Health Rep*, 1962, *77* 503

Phillips, W, Bahn, A K and Miyasaki, M Person Matching by Electronic Methods *Communica ACM*, 1962, *5* 404

Piel, G. Science, Censorship, and the Public Interest *Science*, 1957, *125* 792

Pipberger, H V and Wood, C R A Simplified Method for the Resolution of the Orthogonal Electrocardiogram *Circulat Res*, 1958, *6*,239

Pipberger, H V and Tanenbaum, H L The P-Wave, P-R Interval, and Q-T Ratio of the Normal Orthogonal Electrocardiogram *Circulation*, 1958, *18* 1175

Pipberger, H V The Normal Orthogonal Electrocardiogram and Vectorcardiogram, with a Critique of some Commonly Used Analytic Criteria *Circulation*, 1958, *17* 1102

Pipberger, H V and Lilienfield, L S The Applications of Corrected Electrocardiographic Lead Systems in Man *Amer J Med*, 1958, *25* 539

Pipberger, H V Die Bedeutung der Orthogonalen Elektrokardiographie und Vektorkardiographie fur die Klinische Kardiologie *Arch Kreislaufforsch*, 1958, *29* 58

Pipberger, H V Current Status and Persistent Problems of Electrode Placement and Lead Systems for Vectorcardiography and Electrocardiography and *Prog Cardiov Dis*, 1959, *2* 248

Pipberger, H V, Freis, E D, Taback, L and Mason, H. L Preparation of Electrocardiographic Data for Analysis by Digital Electronic Computer *Circulation*, 1960, *21* 413

Pipberger, H V and Freis, E D Automatic Analysis of Cardiological Analogue Data by Means of Electronic Computers *Med Dokum*, 1960, *4* 58

Pipberger, H V, Bialek, S M, Perloff, J K and Schnaper, H W

Correlation of Clinical Information in the Standard 12-Lead ECG and in a Corrected Orthogonal 3 Lead ECG *Amer Heart J*, 1961, *61* 34

Pipberger, H V, Arms, R J and Stallman, F W Automatic Screening of Normal and Abnormal Electrocardiograms by Means of a Digital Electronic Computer *Proc Soc Exp Biol Med*, 1961, *106* 130

Pipberger, H V and Carter, T N Analysis of the Normal and Abnormal Vectorcardiogram in Its Own Reference Frame *Circulation*, 1962, *25* 827

Pipberger, H V Use of Computers in Interpretation of Electrocardiograms *Circulat Res*, *1962*, *Part 2*, *11* 555

Plackett, R L *Principles of Regression Analysis* Oxford Univ Press, London, 1960

Pratt, A W, Toal, J N and Sober, H A A Computational Approach to the Spectral Analysis of Oligonucleotides *Fed Proc*, *1962*, *21* *376*

Preston, J S Computer—or Tables? *Nature*, 1962, *193* 207

Pronko, E, Donald N R and Smits, R. M Current Research and Development in Scientific Documentation, No 9 *Nat Sci Foundation*, NSF-61-76, 1961

Radcliff, G Automation Versus Socialization in the General Hospital *J Int Coll Surg* 1961, *36* 542

Radin, N A Punched Card System for the Laboratory *Clin Chem*, 1962, *8* 538

Ralston, A and Wilf, H S *Mathematical Methods for Digital Computers* J Wiley and Sons, New York, 1960

Ramo, S Extending Mans Intellect by Electronics *Proc IRE*, 1962, *50* 640

Reiss, R F The Digital Stimulation of Neuro-Muscular Organisms *Behav Sci*, 1960, *5* 343

Revans, R. W *The Hospital As An Organization—A Study in Communication and Morale* Pergamon Press, New York, 1960

Richman, A and Law, A G The Application of Electronic Average Response Computers to Psychiatric Research *Canada Psychiat Assn J* 1961, *6* 202

Riehl, J L Analog Analysis of EEG Activity *Aerospace Med* 1961, 32 1101

Rikli, A E, Tolles, W E, Steinberg, C A, Carbery, W J, Freiman, A H and Abraham, S Computer Analysis of Electrocardiographic Measurements *Circulation*, 1961, *24* 643

Rikli, A E and Caceres, C A The Use of Computers by Physicians as a Diagnostic Aid *Ann N Y, Acad Sci*, 1961, *23* 237

Riordan, J *Stochastic Service Systems* J Wiley and Sons, New York, 1962

Robertson, J S The Use of Analog Computers in the Kinetic Analysis of Compartmental Systems *Trans N Y Acad Sci*, 1961, *23* 506

Robinson, G D *Vector Geometry* Allyn and Bacon, Boston, 1962

Rockoff, M L Diagnostic Aspects of Computer Applications in Medical Research in the University of Pennsylvania *IRE Trans on Med Electronics*, 1960, ME-7, 250

Rockwell, T H, Barnum, R A and Giffen, W C Inventory Analysis as Applied to Hospital Whole Blood Supply and Demand *J Ind Engin*, 1962, *13* 109

Rogers, D J and Tanimoto, T T A Computer Program for Classifying Plants *Science*, 1960, *132* 1115

Rogers, F A The Use of IBM Tabulating Methods in the Analysis of Medical Data *Minnesota Med*, 1961, *44* 382

Roman, J A, Ware, R W, Adams, R M, Warren, B H and Kahn, A R School of Aerospace Medicine Physiological Studies in High Performance Aircraft *Aerospace Med*, 1962, *33* 412

Rome, H P Symposium on Automation Techniques in Personality Assessment *Proc Mayo Clin*, 1962, *37* 61

Roof, S D and Frazier, G D An IBM 709 Computer Program for Computing Bartlett's Test of Homogeneity of Variance *Biometrics*, 1962, *18* 252

Rosenblith, W A Relations Between Auditory Psychophysics and Auditory Electrophysiology *Trans N Y Acad Sci*, 1957, *19* 650

Rosenblith, W A Information and Control in Organ Systems *IRE Trans on Med Electronics*, 1959, ME-6, 23

Royce, J R. Search for Meaning *Amer Scient*, 1959, *47* 515

Rudin, S A Measuring the Teacher's Effectiveness as a Lecturer *J Genet Psychol*, 1961, *98* 147

Rushmer, R. F Physiological Application of Computers *Circulat Res* 1962, Part 2, *11* 529

Russe, O Rational Organization and Documentation in a Large Acci dent Hospital *Med Dokum*, 1960, *4* 27

Rutstein, D D, Eden, M and Schützenberger, M P Report on Mathematics in the Medical Sciences *New Engl J Med*, 1961 *265* 172

Saaty, T L *Elements of Queuing Theory* McGraw-Hill, New York, 1961

Saiger, G L Observations on the Probability of Error in Medical Diagnosis *Ann Int Med*, 1962, *56* 860

Sandel, T T and Kiang, N Y "Off" Responses in the Auditory Nervous System of Anesthetized Cats *J Acoust Soc Am*, 1957, *29* 184

Saunders, M G Digital Conversion of Electroencephalograph Records, *Electronics*, 1960, *33* 78

Scher, A M , Young, A C and Meredith, W M Factor Analysis of the Electrocardiogram Test of Electrocardiographic Theory Normal Hearts *Circulat Res* , 1960, *8* 519

Scher, A M Servoanalysis of Blood Pressure Control Mechanism *15th Annl Conf on Eng In Med and Biology*, 1962, p 53

Schenthal, J E , Sweeney, J W and Nettleton, W Clinical Application of Large-Scale Electronic Data Processing Apparatus *J A M A* , 1960, *173* 90

Schenthal, J E , Sweeney, J W and Nettleton, W Clinical Application of Electronic Data Processing Apparatus *J A M A* , 1961, *178* 267

Szheibe, O On Final Evaluation of Patient Records Before the Period of Machine Documentation *Langenbeck Arch Klin Chir* , 1961, *298* 661

Schmitt, O H Cathode-Ray Tube Presentation of 3-Dimensional Data, *J Appl Physiol*, 1947, *18* 319

Schmitt, O H Application of Computers in Cardiovascular Disease *Circulat Res* , 1962, Part 2, *11* 505

Schmitt, O H Adaptive Analog Models for Biological Rhythms *Ann N Y Acad Sci* , 1962, *98* 846

Schreiber, F and Nielsen, A Punch Card Code for Classification of Craniocerebral Injuries *J Mich Med Soc* , 1938, *37* 909

Scott, R C , Seiwert, V J , Simon, D L and McGuire, J Left Ventricular Hypertrophy A Study of the Accuracy of Current Electrocardiographic Criteria When Compared with Autopsy Findings in one Hundred Cases *Circulation*, 1955, *11* 89

Segall, H N The Electrocardiogram and Its Interpretation A Study of Reports by 20 Physicians on a Set of 100 Electrocardiograms *Canada Med Assoc J* , *1960*, *82* 2

Seibel, R Computer Solutions to Some Noncomputational Psychological Problems *Educ and Psychol MSRMT* , 1961, *21* 185

Sekel, J P , Eng D , Jegier, W and Johnson, A L Automatic Electronic Computer for the Estimation of Arterial Concentration of Evans Blue Dye *Amer Heart J* , 1958, *55* 485

Selfridge, O and Neisser, U Pattern Recognition by Machine *Sci Amer* , 1960, *203* 60

Seligson, D Automation and Labor Saving Devices in the Clinical Laboratory *Conn Med* , 1961, *25* 423

Seligson, D Men and Machines *Clin Chem* , 1962, *8* 203

Shepard, R B Analog Computer Use for Immediate Determination of

Phasic Power and Mean Energy in Arterial Flow *Circulation*, 1961, *24* 1040

Sherer, H Here Is an Example of "Scrambled" Programmed Instruction for Administrators *Mod Hosp*, 1962, *99* 81

Shipton, H W An Electronic Trigger Circuit as an Aid to Neurophysiological Research *J Brit Instn Radio Engrs*, 1949, *9* 374

Shore, M S and Callahan, R. Application of Digital Computer to Study of Dynamic Biological Systems *Fed Proc*, 1962, *21* 292

Siegel, I H Least Squares "Without Normal Equations" *Trans N Y Acad Sci*, 1962, *24* 362

Siler, W and Laughlin, J S A Computer Method for Radiation Treatment Planning Communica ACM, 1962, *5* 407

Silink, K The Possibility of Designing Machines which Learn Diagnostic The Zero-Systems of Types of Pathotypes in Endocrinology *Activitas Nervosa Supr*, 1961, *3* 148

Silverman, M and Burgen, A S V Application of Analogue Computer to Measurement of Intestinal Absorption Rates with Tracers, *J Appl Physiol*, 1961, *16* 911

Simonson, E *Differentiation Between Normal and Abnormal in Electrocardiography* Mosby Co, St Louis, 1961

Singer, J R Electronic Analog of the Human Recognition System *J Amer Optical Soc*, 1961, *51* 61

Slavin, M Automation and the Hospital Pharmacist *Amer J Hosp Pharm*, 1962, *19* 274

Slee, V N The Normal Tissue Rate *Bull Amer Coll Surg*, 1962, *47* 108

Slee, V N Automation in the Management of Hospital Records *Circulat Res*, 1962, Part 2, *11* 637

Sloan, R D Roentgenologic Evaluation of Mesenteric Small Intestinal Obstruction A Statistical Analysis *Amer J Roentgenol*, 1959, *82* 978

Smith, E C Data Processing Considerations in the Handling of Medical Data *IRE Trans on Med Electronics*, 1960 ME-7, 379

Smith, H, Gnanadesikan, and Hughes, J B Multivariate Analysis of Variance *Biometrics*, 1962, *18* 22

Smith, J M, Miller, R A, Marienfeld, C, Hahneman, B and Willard, J Use of Tape Recorded Heart Sounds in Screening of Children for Heart Disease *Circulation*, 1959, *20* 887

Smith, M D Communication with Computers, *Circulat Res*, 1962, Part 2, *11* 501

Smith, P A Some Problems and Approaches to Automation of Medical Diagnosis *Behav Sci*, 1961, *6* 88

Smith, P A, Brouha, L and James, E W Analog Simulation of the

Physiological Responses of Men Working in Hot Environments *Fed Proc*, 1962, *21* 223

Smith, P E Analog Simulation of the Physiological Responses of Men Working in Hot Environments *Proc San Dieto Symp Bio Med Eng*, 1962, 117

Southern, W A Mechanized Processing and Retrieval of Bio-Medical Information *Meth Infor in Med*, 1962, *1* *16*

Spencer, W A and Vallbona, C Digitation of Clinical and Research Data in Serial Evaluation of Disease Processes *IRE Trans on Med Electronics*, 1960, ME-7, 296

Stacy, R. W and Giles, F M Computer Analysis of Arterial Properties, *Circulat Res*, 1959, *7* 1031

Stacy, R. W *Biological and Medical Electronics* McGraw-Hill, New York, 1960

Stacy, R. W Diagnosis of Arterial Disease with Analog Computers *IRE Trans on Med Electronics*, 1960, ME-7, 269

Stacy, R. W Computers in Arterial Research *Circulat Res*, 1962, Part 2, *11* 535

Stahl, W R. Similarity and Dimensional Methods in Biology *Science*, 1962, *137* 205

Stallman, F W and Pipberger, H V Automatic Recognition of Electrocardiographic Waves by Digital Computer, *Circulat Res*, 1961, *9* 1138

Stanley, P E A New Approach to Electrocardiography *15th Annl Conf on Eng In Med and Biology*, 1962, p 49

Stark, L Stability, Oscillations, and Noise in the Human Pupil Servomechanism *Proc IRE*, 1959, *47* 1925

Stark, L, Okajima, M and Whipple, G H Computer Pattern Recognition Techniques Electrocardiographic Diagnosis *Communica ACM*, 1962, *5* 527

Sterling, L F A Factorial Analysis of Cardiovascular Variables *Dissert Abstr*, 1960, *20* 4582

Sterling, T D, Perry, H and Bahr, K. S A Practical Procedure for Automating Radiation Treatment Planning *Brit J Radiol*, 1961, *34* 726

Sterling, T D, Perry, H and Weinkam, F F Automation of Radiation Treatment Planning, II, Calculation of Non Convergent Field Dose Distributions *Brit J Radiol*, 1963, *36* 63

Steward, C B and Manning, G W A Detailed Analysis of Five Hundred RCAF Aircrew *Amer Heart J*, 1944, *27* 502

Stokes, J and Dawber, T R, Rheumatic Heart Disease in the Framingham Study *New Engl J Med*, 1956, *225* 1228

Stokes, J and Dawber, T R The Silent Coronary The Frequency and Clinical Characteristics of Unrecognized Myocardial Infraction in the Framingham Study *Ann Int Med*, 1959, *50* 1359

Stolwijk, J A, Jr and Hardy, J D Theoretical Study of Temperature Regulation in Man *Fed Proc*, 1962, *21* 218

Stoll, R R *Linear Algebra and Matrix Theory* McGraw-Hill, New York, 1952

Storm, V W and Magnus, O The Significance of Automatic Analysis of the electroencephalogram for Clinical Diagnosis *Rev Neurol*, 1961, *104* 177

Strehler, B L and Mildvan, A S General Theory of Mortality and Aging A Stochastic Model Relates Observations *Science*, *1960*, *132* 14

Swenson, W M The Test Symposium on Automation Technics in Personality Assessment *Proc Mayo Clin*, *1962*, *37* 61

Swisher, S and Agalides, E Probabilistic Logic and the Algebra of Functors Applied to Data Disease Complexes *15th Annl Conf on Eng in Med and Biology*, 1962, p 13

Taback, L, Marden, E, Mason, H L and Pipberger, H V Digital Recording of Electrocardiographic Data for Analysis by Digital Computer *IRE Trans on Med Electronics*, 1959, ME-6, 167

Takacs, L *Introduction to the Theory of Queues* Oxford Univ Press, London, 1962

Taylor, J D and Wiegand, R G The Analog Computer and Plasma Drug Kinetics *Clin Pharmacol Ther*, 1962, *3* 464

Taylor, W K Computers and the Nervous System *Sympos Soc Exp Biol*, 1960, *14* 152

Teorell, T Application of the Voltage Clamp to the Electrohydraulic Nerve Analog *Acta Soc Med Upsaliensis*, 1960, *65* 231

Thomas, A J, Cochrane, A L and Higgins, I T T Measurements of Prevalence of Ischaemic Heart Disease *Lancet*, 1958, *2* 540

Thompson, E M and Tucker, H Computers in Dietary Studies *J Amer Diet Assoc*, 1962, *40* 308

Thoma, J A Simple and Rapid Method for the Coding of Punched Cards *Science*, 1962, *137* 278

Thompson, J B, Avant, O W and Spiker, E D How Queuing Theory Works for the Hospital *Mod Hosp*, 1960, *94* 75

Thompson, S W *A Histopathologic and Histologic Coding System for Tissue Collections* Thomas, Springfield, 1961

Thorner, R M and Packer, H Combined Record and Data Processing

System in a Glaucoma Clinic Program *Public Health Rep*, 1961, *76* 1107

Thornton, W , Gibson, T W and Davis, D A Computer Analysis of the Electrocartiogram *15th Annal Conf on Eng In Med and Biology*, 1962, p 48

Tobach, E , Schneirla, T C Aronson, L R and Laupheimer, R. The ATSL An Observer to Computer System for a Multivariate Ap proach to Behavioral Study *Nature*, 1962, *194* 257

Tolles, W E , Carbery, W J , Freiman, A H , Steinberg, C A and Caceres, C A Role of the Modern Computer in Analyzing Electrocardiographic Data *Circulation*, 1960, *22* 824

Tolles, W E , Steinberg, C A , Carbery, W J and Freiman, A H Experimental Techniques and Results of a Study Using a Computer as a Diagnostic Aid *Trans N Y Acad Sci* , 1961, *23* 246

Tsien, K C The Application of Automatic Computing Machines to Radiation Treatment Planning *Brit J Radiol* , 1955, *29* 432

Tukey, J W One Degree of Freedom for Non-Additivity *Biometrics*, 1949, *5* 232

Ueda, H , Ilo, M , Kameda, H and Migita, T Analysis of Radioactive Colloidal Gold Uptake Rate in the Liver by Electronic Digital Com puter *Jap Heart, J* , 1961, *2* 460

Ueda, H , Ilo, M , Nakanishi, A and Koide, T Digital Computer Aids for the Cardiac Output Calculation from Radio Isotope Dilution Curves *Jap Heart J* , 1962, *3* 24

Uhr, L and Vossler, C Suggestions for Self Adapting Computer Models of Brain Functions *Behav Sci* , 1961, *6* 91

Unger, S H Pattern Detection and Recognition *Proc IRE* , 1959, *47* 1737

Uttal, W R and Roland, P A A Terminal Device for Entry of Neuro electric Data into an Electronic Data Processing Machine *Electro enceph Clin Neurophysiol* , 1961, *13* 637

Vanderpol, B *Basic Principles in Vector Electrocardiography* Thomas Springfield, 1961

Vandermark, F An Automatic Electronic Average Response Computer for Physiological Purposes (Evoked Response Detector) *Acta Physiol Pharmacol Neerl* , 1961, *10* 29

Vandenberg, J An Electrical Analog of the Trachea, Lungs, and Tissues *Acta Physiol Pharmacol Neerl* , 1960, *9* 361

Vanwoerkom, A J Program for a Diagnostic Model *IRE Trans on Med Electronics*, 1960, ME-7, 220

Vanwoerkom, A J and Brodman, K Statistics for a Diagnostic Model *Biometrics*, 1961, *17* 299

Vishevskii, A A, Bychovskif, M L and Charnas, S S On the Possible Use of Computing Machines for Diagnostic Purposes *Cas Lek Cesk*, *1961*, *100 385*

Vonfoerster, H Communication Amongst Automata *Amer J Psychiat*, *1962*, *118 865*

Voskresendkij, A D and Prokhorov The Use of Electronic Computers in Various Fields of Medicine *Trans Sovet Zdravookhr*, *1959*, *18* No 8, 19

Wade, L Electronic Data Processing A Scheme for an Occupational Medical Service *A M A Arch Environ Health*, 1961, *3* 526

Wajchenberg, B L, Pieroni, R R and Hoxter, G Preliminary Mathematical Model for Glucagon-Induced Hepatic Glycogenolysis in Man *Metabolism*, 1961, *10* 585

Walker, A E, Henry, C E and Merlis, J K A Symposium Recent Developments in Electroencephalographic Techniques, Third International Congress of Electroencephalography and Clinical Neurophysiology *Electroenceph Clin Neurophysiol*, 1953, Suppl 4

Walker, L A Frontiers of Medicine, 11, Electronic Computers in Medicine *Northw Med*, 1960, *59* 658

Walsh, J E Analyzing Medical Data—Some Statistical Considerations *IRE Trans on Med Electronics*, 1960, ME-7, 362

Walsh, J E *Handbook of Nonparametric Statistics Investigation of Randomness, Moments, Percentiles, and Distributions* Vannostrand Co, New Jersey, 1962

Ward, J H Markov Models and Monte Carlo Techniques *Educ and Psychol Msrmt*, 1961, *21* 219

Ward, J H and Hook, M E Use of Regression Analysis and Electronic Computers in the Prediction of Coronary Artery Disease *Behav Sci*, 1962, *7* 120

Warfield, J N *Introduction to Electronic Analog Computers* Prentice Hall, Englewood Cliffs, 1959

Warner, H R. The Frequency Dependent Nature of Blood Pressure Regulation by the Carotid Sinus—Studies With an Electronic Analog *Circulat Res*, 1958, *6* 35

Warner, H R. The use of an Analog Computer for Analysis of Control Mechanisms in the Circulation *Proc IRE*, 1959, *47* 1913

Warner, H R , Veasey, L G , Toronto, A F and Stephenson, R E , Diagnosis of Congenital Heart Disease from Clinical Data with a Digital Computer *Circulation*, 1960, *22* 829

Warner, H R and Toronto, A F Regulation of Cardiac Output Through Stroke Volume *Circulat Res* , 1960, *8* 549

Warner, H R , Toronto, A F , Veasey, L G and Stephenson, R. A Mathematical Approach to Medical Diagnosis Applications to Congenital Heart Disease *J A M A* , 1961, *177* 177

Warner, H R Use of Analog Computers in the Study of Control Mechanisms in the Circulation *Fed Proc* , 1962, *21* 87

Warner, H R. and Cox, A A Mathematical Model of Heart Rate Control by Sympathetic and Vagus Efferent Information *J Appl Physiol*, 1962, *17* 349

Watt, I E F Use of Mathematics in Population Ecology *Ann Rev Entomol*, 1962, 7 243

Weaver, W Science and Complexity *Amer Scient* , 1948, *36* 536

Weinrauch, H and Hetherington, A W Computers in Medicine and Biology *J A M A* , 1959, *169* 240

Weiss, I M and Hoover, K. H *Scientific Foundations of Education* Wm C Brown, Co , Dubuque, 1960

Wertz, F E Electronic Data Processing for Hospitals—Current Practices and Future Needs *Hospitals*, 1962, *36* 44

Wessel, H U , Jacobs, J E , Despe, O D and Kezdi, P Evaluation of a Simple Computer for Cardiac Output Measurements Based on Thermal Dilution *IRE Trans on Bio-Med Electronics*, 1962, BME-9, 155

White, B W Computer Applications to Psychological Research Studies in Perception *Behav Sci* , 1962 *7* 396

Whitney, D R. *Elements of Mathematical Statistics* H Holt & Co , New York, 1959

Wiener, N *Cybernetics* J Wiley and Sons, New York, 1948

Wilcox, P H Electrostimulation Therapy and Drugs A Future for Computers *Dis Nerv Syst* , 1961, *22* 50

Williams, E J *Regression Analysis* J Wiley and Sons, New York, 1959

Williams, J A and Fine, J Measurement of Blood Volume with a new Apparatus *New England J Med* 1961, *264* 842

Williams, P Note on the Sampling Distribution of β_1, *Where the Population is Normal* *Biometrika*, 1935, *27* 269

Woodbury, M A and Lipkin, M Coding of Medical Case History Data for Computer Analysis *Communica ACM* , 1962, *5* 532

Woodbury, M A What Are Computers? *Circulat Res* , 1962, PART 2 *11* 485

Wullstein, H L , Bandtlow, O , Kreutle, O J and Schmitt, H G Investigations of the Statistical Evaluation of Operations for Improvement of Hearing with IBM Punch Card Methods *Acta Otolaryng* , 1962, *54* 56

Yanof, H and Dobson, E L Analog Model to Illustrate a Homeostatic Mechanism Accounting for the Physiological Regulation of Body Fluid Volume *U S A E C Rept* , UCRL-8705, 1959

Yano, K and Pipberger, H V Recognition of Left Ventricular Overload Patterns in the Orthogonal ECG and VCG *Clin Res* , 1962, *10* 184

Yates, F Marginal Percentages in Multiway Tables of Quantal Data with Disproportionate Frequencies *Biometrics*, 1961, *17* 1

Yeh, G C K and Martinek, J Multipole Representation of an Eccentric Dipole and an Eccentric Double Layer *Bull Math Biophys* , 1959, *21* 33

Young, L R Discontinuous Biological Control—The Eye Movement System *15th Annl Conf on Eng in Med and Biology*, 1962, p 32

Young, T Y and Huggins, W H Signal Analysis by a Digital Computer Using Orthonormal Exponentials *15th Annal Conf on Eng In Med and Biology*, 1962, p 48

Young, T Y and Huggins, W H On the Statistical Analysis of Electrocardiograms *15th Annl Conf on Eng in Med and Biology*, 1962, p 49

Zimmer, H Preparing Psychophysiologic Analog Information for the Digital Computer *Behav Sci* , 1961, 6 161

Zimmer, H Digital Computer Analysis of Psychophysiologic Information *15th Annl Conf on Eng in Med and Biology*, 1962, p 35

Zusman, F S and Ledley, R S Biomedical Information Retrieval by Means of Tabledex

PART 1 Bibliography Words for Retrieval

PART II Alph Descriptive Words for Part I

PART III Indexing Table

15th Annl Conf on Eng in Med and Biology, 1962, p 22

AUTHOR INDEX*

Abildskov, J A., p 68, L 22
Acheson, R M., p 73, L 23, L. 26
Agalides, E , p 44, L. 16
Allen, L.E , p 122, L. 4

Barlow, J S , p 56, L 7
Bloch, T , p 77, L. 4
Blumberg, M.S , p 34, L 27
Bramson, M.L , p 69, L 29
Brodman, K., p 46, L 7
Brody, D A., p 60, L 30
Bush, V , p 37, L. 29

Caceres, C.A., p 63, L 4, L 7, L 20
Chance, B , p 82, L 25
Clynes, M , p 80, L 26, p 41, L 35
Cox, A., p 82, L. 4

Daniel, C.M , p 63, L 5
Davies, L G , p 61, L 17
Dawber, T R., p 73, L. 32, p 76, L 4
Dayhoff, M D , p 77, L. 12, p 78, L 6
Dewaart, A., p 58, L 27
Doyle, J T , p 76, L. 4

Einthoven, W , p 58, L 27, p 79, L. 24
p 59, L 18
Enslein, K., p 67, L 6
Epstein, F H , p 61, L. 3
Erdmann, A.J , p 46, L.7
Ezekiel, M , p 109, L 9

Fahr, G , p 58, L. 27, p 59, L. 20
Flagle, C.D , p 34, L. 27
Fox, K A., p 109, L 9
Frank, E., p 64, L 7

Gerbarg, D.S , p 49, L. 5
Gerbode, F , p 69, L. 29
Goldstein, H , p 46, L 7
Gordon, T , p 73, L. 32
Gouras, P , p 57, L 14
Gustafson, J E., p 63, L 5, L. 22
Hahneman, B , p 48, L. 38
Hardy, J D , v 38 -p 27
Harken, D , p 69, L. 18
Heslin, A. S , p 76, L. 4
Hidalgo, J U , p. 77, L. 4
Hisey, B L , p 68, L 22
Huggins, W H , p 66, L. 28, p 67, L 10

Ingerson, W E , p 68 L 22

Janaske, P C , p 121, L 31
Kahn, H A , p. 76, L 4
Kane, S H , p 33, 19
Kannel, W.B , p 76, L 4

Ledley, R.S , p 77, L. 12, p 78, L. 6,
p 125, L. 13
Lefemine, A.A., p 69, L. 18
Lindquist, E.F , p 27, L 5, p 90, L 6
Lipkin M , p 38, L 27, p 39, L 33
Luhn, H P , p 120, L. 20
Lusted, L B , p 43, L. 35, p 44, L. 2

Marienfeld, C., p 48, L. 38
Meredith, W M., p 63, L 34
Miller, R.A., p 48, L 37
Moore, F E , p 73, L. 32

---

*See footnote, p 135

Nadler, S B , p 77, L 4
Nash, F.A , p 38, L 5
Newell, A , p 41, L. 53, p 42, L. 8

Osburn, J O , p 69, L 29
Piel, G , p 122, L 12
Pipberger, H V , p 62, L. 11, p 63, L. 26, p 65, L. 18, p 66, L. 8

Romans, W E , p 60, L 30
Russi, M , p 69, L 29

Sale, A , p 69, L. 29
Scher, A.M., p 63, L 34, p 64, L. 7
Schmitt, O H , p 68, L. 14, p 80, L. 10
Segall, H N , p 61, L 9
Shipton, H W , p 57, L. 17, p 58, L. 17
Shurtleff, D , p 73, L. 32
Simon, H A., p 41, L. 33, p 42, L. 8
Slee, V N , p 29, L 34
Smith, J M., p 48, L. 37
Spiker, E.D , p 90, L. 9
Staples, L.F , p 63, L. 6
Stark, L , p 80, L. 34
Steinberg, C.A , p 63, L 5
Stephenson, R , p 40, L. 6
Swenson, W M., p 46, L 27
Swisher, S , p 44, L. 16

Toronto, A.F , p 40, L. 6

Vanwoerkom, A J , p 46, L. 7
Veasey, L.G , p 40, L. 6

Waller, A D , p 58, L. 24
Warner, H R., p 40, L. 6, p 80, L. 18, p 82, L. 4
Wiener, N , p 54, L. 27
White, W J , p 63, L. 5
Willard, J , p 48, L. 38

Young, A.C., p 63, L. 34
Young, T Y , p 66, L. 28, p 67, L. 10

Zimmer, H , p 79, L. 1
Zusman, F S , p 125, L. 13

SUBJECT INDEX*

Additivity in the analysis of variance, p 99, L 8
Administrative significance of computerized medical studies, p 34
Administrative uses of a computer, p 20, 21, 22, 23, 28
Algebra of logic, p 44
American Institute of Biological Sciences, p 121, L 29
American Medical Association, p 34, L 9
American Medical Research Foundation Paranatal Study, p 33, L 15, p 34
Analog computer simulation, p 4-5 inc , p 7
Analysis of control of eye movements, p 55, 56
Analysis of phonocardiogram by a digital computer, pp 48 49
Analysis of variance p 90 L 4, p 92, L 12, pp, 104, 105, 106
Anesthesiology, pp 56-57
Automated medical record systems, p 22, L 10 13, p 25, 48
Automation and labor-saving devices, p 3, 129, 130
Automation of billing procedures, p 25,
Automation of medical diagnosis, p 3
Automation in the general hospital, p 28, 29, 30

Baltimore Accounting Study, p 25, L. 27 33
B.A.S I C., p 124, L. 32
Beta coefficients, p 115, L 3
Biological abstracts p 120, L. 13
Block diagram, p 95, L. 26

Care, organization for hospital inpatients, pp 21 22
Cataloging knowledge, p 121, 122
Changing concepts in patient care, p 30
Chemistry abstracts p 120, L 12
Chi square, p 104, L 5
Cigarette smoking and coronary heart disease, p 76
Clinical research, Storage and retrieval of the results, p 15, L 12 22
Coding of information p 33, 126, 127, *128*
Coding of medical case history data for computer analysis, p 21, 22 31, L 1
Coefficients of regression and determination p 109, L 33
Commission on Professional Hospital Activities Inc., p 28, L 16
Committee on Maternal and Child Care of the Council on Medical Service, p 34, L 12
Computer centers, p 26, 27, 32
Computer logic, p 12
Computer program, p 4
Computer science, p 27
Computers defined, p 4, 8
Congenital heart disease, Computer diagnosis, p 40
Cornell Medical Index, p 45, L *28*, p 46
Correlation analysis, p 18, L 19 29
Cost of data processing p 14, L. 11
Covariance, p 115, L. 26
Current Contents, p 119, L 28
Current research and development in scientific documentation, p 121
Cybernetics, p 54

Data processing apparatus, p 14, 15, 25, L. 3
Data processing for hospitals, p 22

---

*See footnote, p 135

Data processing of patient's records, p 44-45
Data reduction, p 53
Design and analysis of experiments, p 90, L 8
Diagnosis by computer, p 37
Diagnosis of cardiovascular disease, p 63, L 4
Diagnosis of ventricular hypertrophy, p 63, L 5
Diagnosis, Mathematical models in medicine, p 44, 45
Differential diagnosis of hematologic diseases aided by mechanical correlation of data, p 38, 39
Differential diagnosis, An apparatus to assist the logical facilities, p 38, 40, 41, 42, 43, 44-46
Digital and analogue computers, p 4-6, 8, 9
Digital computer simulation, p 4, 5, 8, 9
Digital computers and medical logic, p 7
Dynamic simulation, p 71, 72, 80, 81

Electrocardiogram, p 50, 58, 59 69 inc
Electroencephalogram, p 50, 52, 54, 55, 56, 58
Electroencephalographic techniques, p 51-52
Electromotive forces of the heart, p 60
Error and information, p 13, L 14-21, p 18, L 3
Experiment design and analysis, p 16, 17, 18
Eye movements, p 55 56, 57

Factor analysis, p 44, L 26
Factor analysis of the electrocardiogram, p. 63, L 34
Flow chart, p 12, L 13-23, p 95, L 37
Format cards, p 99, L. 26
FORTRAN, p 98
Framingham Study, p 72-76 inc
F-ratio, p 103, p 104, L 37

General Problem Solving Program (G PS), p *42, L 10 11*, p 42 43
Heart sounds in screening of children, pp 48-49
Human heart, p 58, 59
Hospital
  automation, p 22, L 16
  facilities, p 21, L. 14
  inpatient care, p 21, L 31
  planning, p 21, L 24
  records, p 29, L 26
  research and administration, p 10 11, 20
  Hospitals current practices, p 31, *L 23*
Hospitals—When should they install computers?, p *34, L* 23
Hypotheses, p 15, L 28

Index Medicus p 120, L. 12
Information handling and science in *formation, p 121 L 31*
Information in electrocardiac signal, p 66, 12
Information retrieval, p 14, L 31
Information in electrocardiac signal p 66, L 12
Information retrieval, p 14, L 31
Information specialist, p 128, L 28
Information storage, p 119 126 inc
Interpretation of analyses, p 105, L 34
Interpretation of electrocardiograms, p 62
Iowa Medical Service Plan, p 26, L 30, p 27
Iowa Testing Program, p 27, L 5

Keyword in-context index (KWIC), p 120, 121
Kurtosis, p 95, L 11

Library programs, p *99, L 34*
Logic, Computers and medical, pp 42-40
Logic theory machine, p 41, L. 33 34, p 42

Machine language, p 8, 12, 13, L 18
Mark sense cards, p 25, L 10, p 25
Medical Audit Program (MAP), p 29, *L 10*

Medical practice Some comments on electronic data processing, pp 31-32
Minnesota Multiphasic Personality Inventory, p 45, L. 29, p 46, 47
Moment, Ratios of normal samples, p 95, L 27
Multi font optical character reader, p 67
Multiple regression analyses, p 44, p 108, L. 7

National Library of Medicine, p 126, L 16
National Science Foundation, p 121, L. 7
Nebraska Act, p 34, L 12, p 35
Noise in problem solving, p 63
Normality tests, p 95, L 21
Nursing, Effective use of personnel, p 21, L 5 11

Observer variation, p -61, L 17
Observer error, p 61, L 3
On line computation, p 6, 14, L. 6, p 32, L. 14, p 45, L. 14, p. 71, 78

Patient data, A computer based system, p 23, L 23-30
Pattern-recognition, p 67-68
Personnel, p 27
Personality assessment, p 46
Planimeter, p 5
Probability theory, p 44
Problems and approaches to automation of medical diagnosis, p 10
Professional Activities Study (PAS), p 28, L 34, p 29, L 16
Programming, p 2-13
Protein biochemistry, p 76, 77, 78

Regression analysis, p 76-77, 107, 108, 109, 110, 113
Respiraoory heart rate control, p 41, L 35-37
Results of a study using a computer, p 18

Rhythm abnormalities, p 52, 62

Scientific information notes, p 121, L. 9
Servomechanism in the human pupil, p 80, 81
Signal detection in noise, p 16, 17
Simulation in biochemistry, p 81, 82, 83
Simulation of a hospital on a computer, p 24, 25
Skewness, p 95, L 10
State University of Iowa Computer Center, p 26-28, p 99 101
Statistical adjustment of data, p 16 17
Statistical methods, p 11, L 12, p 52, 53
Statistical models, p 52, 53, Chap 6
Statistics, Gaussian, p 53, L 18
Storage and retrieval, p 8, 14, 15
Systems Development Corporation, p 23, L 23

Tabledex, p 125, L 14
Tape-recorder heart sounds, p 48-49
Teaching aid, p 40, L 34
Telecardiogramme, p 59
Theory of probability, p 44
Thinking Problem solving by computer simulation of thought, p 9-10, 11 12, 17, L 28, p 41, L 33-34
Toposcope, p 57
Transformations in analysis of variance, p 92, L. 32
Transponder function, p 71, 79
Trigger circuit, p 57, 58
Tulane Medical School, p 25, L. 9, p 48
Use of computers, p 13, 14, L 14, p 70

Vector analysis, p 64, 65
Vectorcardiography, p 64, L 8, p 68
Visual function, p 57
Volemetron, p 27, L. 14

Walnut System, p 125, L. 19
Weighting technique, p 39, L. 21
Why automation can't replace the doctor, p 45